2020 Census Detailed Operational Plan for:
10. Paper Data Capture Operation (PDC)

A New Design for the 21st Century

Issued: October 31, 2022
Version: v2.0
Prepared by: Decennial Census Management Division

United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Final

United States Census 2020

[Page intentionally left blank.]

Approvals

This PDC Detailed Operational Plan has been reviewed and approved for use.

Electronically signed	April 14, 2022
Raymond Muenzer IPT Lead	Date Signed
Electronically signed	September 19, 2022
Alexa Jones-Puthoff IPT Program Manager/Assistant Division Chief (ADC)	Date Signed

[Page intentionally left blank.]

Document Change History

Revision #	Version	Date	Description
1	V1.01	8/21/2019	• Initial working DRAFT Version from 2020 Census DOP template, using updated design and published DOP v1.0 as sources. • Uses annotated version of PDC BPM v3.0 dated July 10, 2019 (with cosmetic fixes).
2	V1.02	10/14/2019	IDEF0 updates, cross-DOP consistency IDEF0 updates, minor adjustments to Data Collection IOD, latest updates to BPM diagrams, and various edits.
3	V1.03	11/22/2019	Additional information added to closeout, other minor additions/corrections
4	V1.04	8/17/2020	Operational updates through February 2020
5	V1.05	6/7/2021	Consistency edits, BPM and other updates added
6	V2.0	9/22/2022	Executive Review updates

Note: The fields below control the document version, date, and status in the page footers throughout the document.

Document Footer Information Control Table

Field Name	Version, Date and Status
DocVersion:	Version 2.0
DocDate:	October 31, 2022
DocStatus:	Final

Table of Contents

List of Figures

List of Tables

[Page intentionally left blank.]

1. Document Purpose

The 2020 Census Detailed Operational Plan (DOP) for the Paper Data Capture (PDC) operation is intended for use by U.S. Census Bureau managers, staff, contractors, and other internal and external stakeholders working on the 2020 Census. The document presents the detailed operational design for the 2020 Census PDC operation and includes a summary of the operational processes involved, their inputs, outputs, controls, and the basic mechanisms employed to conduct the operational work.

Anticipated uses of this document include the following:

- Communication—Documents operational design details for internal and external stakeholders.
- Planning—Documents planning assumptions and key milestones.
- Staffing—Documents staffing needs and strategies.
- Design—Describes operations and flows, which inform design of information technology (IT) systems, manual processes, and training.
- Development—Identifies business rules and required capabilities to be developed.
- Testing—Provides a basis for developing integrated test plans for IT systems and processes.

This document complements the 2020 Census Operational Plan, which presents the 2020 Census operational design and covers all operations required to execute the 2020 Census, starting with pre-census address and geographic feature updates, and ending once census data products are disseminated and coverage and quality are measured.

2. Operational Overview

2.1 Operation Purpose

The Paper Data Capture (PDC) operation captures and converts data from 2020 Census paper questionnaires. This operation includes:

- Mail receipt
- Document preparation
- Scanning
- Optical Mark Recognition (OMR)
- Optical Character Recognition (OCR)
- Key From Image (KFI)
- Data delivery
- Checkout
- Form destruction

2.2 Background

For Census 2000 and the 2010 Census, PDC systems development and operations were outsourced, with the exception of operations at the National Processing Center (NPC), which were executed under government leadership (and remain under this type of leadership). After the 2010 Census, the integrated Computer Assisted Data Entry (iCADE) system was introduced as a decennial census paper processing alternative through the Improving Operational Efficiency (IOE) program, and it was assessed to be a viable system to use to conduct PDC during the 2020 Census. The iCADE system is a large-scale, efficient, and accurate data capture system that incorporates both automated and manual data capture functionalities. It is currently in operation for several ongoing censuses and surveys at the U.S. Census Bureau.

After the 2010 Census, the Census Image Retrieval Application (CIRA) was developed to provide Census Bureau analysts secure access to images of questionnaires and the corresponding response data needed for research and analysis. It features a centralized query system providing retrieval of and access to questionnaire images and associated captured data. To provide an archive of paper response images, CIRA and iCADE were integrated. For the 2020 Census, CIRA will display all data visually to support the Age Search Operation.

The PDC operation is responsible for the capture and conversion of data from self-response and personal visit paper questionnaires. For the 2020 Census, questionnaires delivered by the

United States Postal Service (USPS) are processed by predominantly temporary staff NPC hires to work at the Paper Data Capture Center (PDCC) East in Jeffersonville, Indiana, and at PDCC West in Phoenix, Arizona. The Island Areas Censuses (IAC) questionnaires are processed by NPC permanent and temporary staff at NPC's main campus, also located in Jeffersonville, but separate (7 miles away) from the location of PDCC East.

For the 2020 Census, it is expected that the majority of self-responses will be provided online using an Internet Self-Response (ISR) instrument. The 2020 Census plan is to promote the internet as the preferred mode of enumeration and to provide Census Questionnaire Assistance (CQA) as a secondary response option. Maximizing the use of an ISR instrument offers opportunities such as reducing the paper processing footprint of PDC processing facilities. The plan also reduces the number of PDCC sites needed from three to two and includes using the iCADE system to capture paper questionnaire data. The paper questionnaires themselves are stored until verification that the data have been received by headquarters (HQ); the paper questionnaires are subsequently destroyed per security regulations.

The PDC operation is largely driven by the timing of the questionnaire mailout, volume of forms received, timing of the initial Nonresponse Followup workload universe cut, and any necessary priority capture requirements. Data are captured from the paper forms in the most efficient manner possible, and both data and images of the forms are maintained. The data are sent to the Response Processing Operation, and the images are sent to the Archiving operation. Captured data from scanned paper images are maintained at NPC, in CIRA, and are made available to conduct the Age Search service.

The PDC Integrated Project Team (IPT) is using 20.5 million forms as the 2020 Census workload estimate. For capacity planning, PDC uses 31.7 million forms to allow for contingency against the potential receipt of a larger-than-expected volume of paper questionnaires being returned based on a lower-than-expected response rate from the internet and CQA.

Mail returns are identified using USPS postal tracing to indicate that a form is en route to a processing center. Upon receipt at a processing center, mail return questionnaires are processed in First-In-First-Out (FIFO) order, unless otherwise specified. The document preparation area removes mail returns from the envelopes and prepares them for scanning. Booklet forms have the binding (spine) removed via the use of a guillotine.

The questionnaires are then delivered to scanning to begin the data capture process. All questionnaires are scanned by iCADE (there is no key from paper). Once scanned, the questionnaires are physically moved to the checkout area. There, questionnaires are staged

while the operation awaits confirmation that questionnaire data have been successfully transmitted.

Scanned images are sent forward for further processing through the iCADE system in which Optical Mark Recognition (OMR) and Optical Character Recognition (OCR) are performed. Data fields with low confidence OMR fields are sent to Manual Registration. Data fields with low confidence OCR results are sent to the Key From Image (KFI) process. Both data and images are maintained. Data are sent to the Response Processing Operation (RPO) at HQ and images are archived locally at the PDCCs and at the Census Bureau's data center, known as the Bowie Computer Center (BCC), located in Bowie, Maryland. Once data have been received at HQ, questionnaires are checked out to ensure the data from each questionnaire have been captured. Once confirmation is received, questionnaires are then eligible for destruction per security regulations.

For the 2020 Census, operational innovations include:

- Reducing the PDC operational workload and associated infrastructure by using ISR and by automating field operations.
- Using an in-house system (i.e., iCADE) to conduct PDC.
- Using USPS tracing data to identify returned questionnaires prior to their arrival at the processing centers (reducing follow-up workloads for other operations).

The following paragraphs describe how PDC was performed for both Census 2000 and the 2010 Census. A brief discussion of the key innovations implemented in the 2020 Census is also provided.

Census 2000

The Data Capture System 2000 (DCS2000) provided the hardware and software to capture questionnaire data for Census 2000. The DCS2000 captured data through a scanning process to create a digital image that was passed through OMR and OCR devices. Manual keying was used to enter data not captured electronically. The data were then aggregated to support detailed tabulations.

For Census 2000, four data capture centers were used. Three were contracted sites in Baltimore, Maryland; Phoenix, Arizona; and Pomona, California. The fourth site was NPC Main Campus in Jeffersonville.

The DCS2000 efficiently and effectively processed approximately 162 million forms. The Baltimore site processed approximately 39 million; the Phoenix site processed approximately

46 million; and the Pomona site processed approximately 44 million. NPC Main Campus processed approximately 32 million forms.

2010 Census

For the 2010 Census, the Decennial Response Integration System (DRIS) was used for the PDC operation.

The PDC operation consisted of three data capture centers. Two were contracted facilities, in Phoenix, Arizona and Baltimore, Maryland. The third was NPC Main Campus in Jeffersonville, Indiana. The Phoenix and Baltimore locations each completed approximately 40 percent of the paper data capture workload, and NPC completed the remaining 20 percent.

In 2010, the data capture centers processed and captured data from over 164 million paper questionnaires, which contained over 3 billion individual checkbox fields. The DRIS had more than 50 different interfaces (system data exchanges) with various systems at Census Bureau HQ during peak production and captured and converted over 65 form types. In addition to the 164 million questionnaires processed in calendar year 2010, the DRIS also processed more than 2 million forms at NPC for Group Quarters validation operations in 2009.

Based on lessons learned from the 2010 Census studies and reviews, the following recommendations were made:

- Ensure development of a timely and comprehensive forms list to facilitate timely system template builds and testing.
- Assign ownership to every field on a questionnaire for when questions arise.
- Ensure timely and realistic contingency planning to properly estimate the PDC workload.
- Leverage postal tracing to monitor inbound and outbound mailings to inform staffing levels.
- Ensure barcode serialization on forms to retain integrity of separated booklets, allowing the system the unique ability to associate booklets in the unlikely event the pages become separated.

2020 Census

For the 2020 Census, the majority of self-responses are expected to be provided on the internet using an ISR instrument. The plan is for the 2020 Census to promote the internet as the preferred mode of enumeration and to provide a phone response option via CQA.

Most respondents are sent internet invitation letters containing a Census ID encouraging them to respond on the internet. Paper questionnaires are sent to an estimated 20 percent of the self-response area universe and are provided to housing units in the Update Leave (UL) universe. Households in self-response and UL areas are sent reminder cards or letters, or both, asking them to respond. All nonresponders in self-response areas receive a paper questionnaire. Special populations, such as those residing in group quarters (GQ), are enumerated using a variety of methods, including paper data collection.

The anticipated workload for PDC for the 2020 Census is 20.5 million forms, and 31.7 million forms is being used for capacity planning. PDCC East is anticipating a total form count workload of 15.7 million and a total page workload of 103.8 million. PDCC West is anticipating a total form count workload of 16.0 million and a total page workload of 153.1 million. For the Island Areas Censuses, NPC Main Campus is anticipating a total form count of 223,000 and a total page workload of 8.9 million.

For the 2020 Census, the Census Bureau relies on USPS to provide postal tracing data for questionnaires coming to the PDCCs.

2.3 Design Overview

The sections below present the high-level design for the PDC operation. Please refer to the 2020 Census Operational Plan for a complete inventory of design decisions for all 2020 Census operations.

2.3.1 High-Level Operational Design

The design of the PDC operation for the 2020 Census includes 11 major operational activity areas:

- PDC Planning, Preparation, and Testing.
- PDC Universe Management.
- Questionnaire Receipt and Check-in.
- Paper Questionnaire Data Capture.
- PDC Quality Assurance Check.
- Response Translation.
- Data Distribution.
- Paper and Image Disposition.
- Paper Data Quality (PDQ) Quality Assurance.
- Island Areas Censuses PDC.

- PDC Closeout.

Each of these major activity areas is summarized below. Together, these activities represent the complete set of work that needs to be performed to conduct this operation.

PDC Planning, Preparation, and Testing – PDC performs activities needed to plan and prepare for production. This includes testing activities, such as testing conducted using printed samples of questionnaires; system and sorter configuration activities; capacity planning; equipment and facilities buildout; and hiring and training of staff.

PDC Universe Management – The Response Processing Operation (RPO) provides the enumeration universe to PDC prior to the start of mailout (contact) operations. RPO provides updates to the universe as work completes in all modes.

Questionnaire Receipt and Check-in – The PDC operation receives mailed questionnaires from the USPS and stages and sorts them. Forms are separated from envelopes by hand. Forms are then stacked in trays. Prior to the 2018 End-to-End (E2E) Census Test (CT), utilizing an extractor for envelope removal was determined to not be a viable method for the 2020 Census and was descoped.

Paper Questionnaire Data Capture – Paper questionnaires are converted to digital images through the scanner; image processing is conducted; and write-in responses are captured, recognized, and processed via OCR, OMR, and Key From Image (KFI), as necessary.

PDC Quality Assurance Check – Data capture quality output verification is performed on automated and manually produced results. Data quality levels taken from a sample of fields identified on the questionnaire are verified to confirm they are meeting quality assurance (QA) requirements. If they do not meet QA requirements, additional fields within the batch are verified.

Response Translation –This activity is conducted to process language referrals. It is used for Hispanic origin and race write-in fields ONLY.

If a respondent provides an answer in a non-English/non-Spanish language, a keyer flags that response.

If that response is in one of the supported foreign languages, it is then translated by translation staff and keyed as the English equivalent. The translation staff then set the appropriate flag of “translated OK.”

If the response is in one of the unsupported languages or is unintelligible, the field is blanked, and the “translate fail” flag is set.

If the response is identified by a translator to be in English or Spanish and should not have been sent for translation, translation staff key the response as written and set the "translate NA" flag.

Data Distribution – Paradata are sent to the Program Management Operation (PM). Response data and case status are sent to the Response Processing Operation (RPO).

Paper and Image Disposition – At the completion of PDC operations, images of the paper questionnaires are made available to the Archiving operation. Paper questionnaires are destroyed in accordance with security requirements once all Census Bureau data capture and image retention requirements are met.

Paper Data Quality (PDQ) Quality Assurance – Once production ends, PDQ is performed to provide valuable metrics for operational assessments and measurement against Service Level Agreements (SLAs). Moving PDQ to occur after production is due to recruiting and hiring concerns at PDCC East and NPC Main Campus.

Island Areas Censuses (IAC) PDC – The Census Bureau conducts the 2020 IAC through contract agreements with local government agencies in American Samoa, Commonwealth of Northern Mariana Islands (CNMI), Guam, and the U.S. Virgin Islands. The Census Bureau provides the materials and guidance to the local government agencies that are then responsible for recruiting and hiring the staff to conduct the data collection activities. The NPC Main Campus processes all IAC forms using the American Community Survey (ACS) process flow. (Note: this PDC operation is conducted as part of "later operations" and follows a separate schedule and timeline.)

PDC Closeout – The closeout of the 2020 Census PDC operation occurs once all forms received have traversed through the processes described above and all data have been delivered; forms are then ready for destruction. Questionnaires are destroyed once all Census Bureau data capture and image retention requirements are met. To close out and decommission the PDCCs, a complete clean out of all equipment, furniture, and materials is necessary, as well as the release of temporary staff and the return of facilities to the respective property owners. (Some of these closeout activities are not relevant to IAC PDC.)

The full hierarchy of activities for the PDC operation is provided in Appendix C in the form of an Activity Tree. In the Activity Tree, each major operational activity area listed above is numbered and then decomposed into a numbered set of subactivities, some of which are further decomposed into more detailed numbered subactivities or steps.

Changes to the operation since the publication of the v1.0 Detailed Operational Plan

A major change to the PDC operation is the timing of the PDQ activity. PDQ was initially scheduled to be conducted during the production period but was rescheduled to be conducted after the data capture production period is complete.

Another major change to the PDC operation is the end date of the operation. Consistent with the COVID-19 pandemic-related extension of response operations, the PDC operation was extended 90 days in accordance with executive direction.

For a full description of the operational subactivities that comprise the PDC operation, see the Detailed Process Description discussions in Section 3 below.

2.3.2 PDC Operational Context

The PDC operational activities described above are conducted within the context of other 2020 Census operations and other programs or data sources that are external to the 2020 Census Bureau. One way to depict an operational context is by using a "Context Diagram," which shows the boundary of the operational process, the operational activities it contains, and the information exchanged with its neighbor operations (or other entities) as well as the resources (mechanisms) needed to conduct the operational work.

Figure 1 is a top-level context diagram for the PDC operation represented as an Integrated Definition, Level 0 (IDEF0) model. An IDEF0 model of a process (or operation) shows the Inputs, Controls, Outputs, and Mechanisms of the process. These IDEF0 model elements are summarized below and described further in the sections that follow.

The yellow box in the center of the IDEF0 model lists the major operational activity areas for the operation, numbered as given in the PDC operation Activity Tree in Appendix C. Specific Information Exchanges (IE) are shown in different colored boxes to represent the Inputs (green boxes on left side), Outputs (orange boxes on right side), Controls (purple boxes on top), and Mechanisms (blue boxes on the bottom). Boxes to the left of the Inputs indicate the *Provider* of the inputs to the operation (typically another 2020 Census operation or an external source). The Provider of the Controls is noted in the box itself. Boxes to the right of the Outputs indicate the *Receiver* of the outputs (typically another 2020 Census operation or external entity). Each Information Exchange has a name and a unique number for identification purposes.

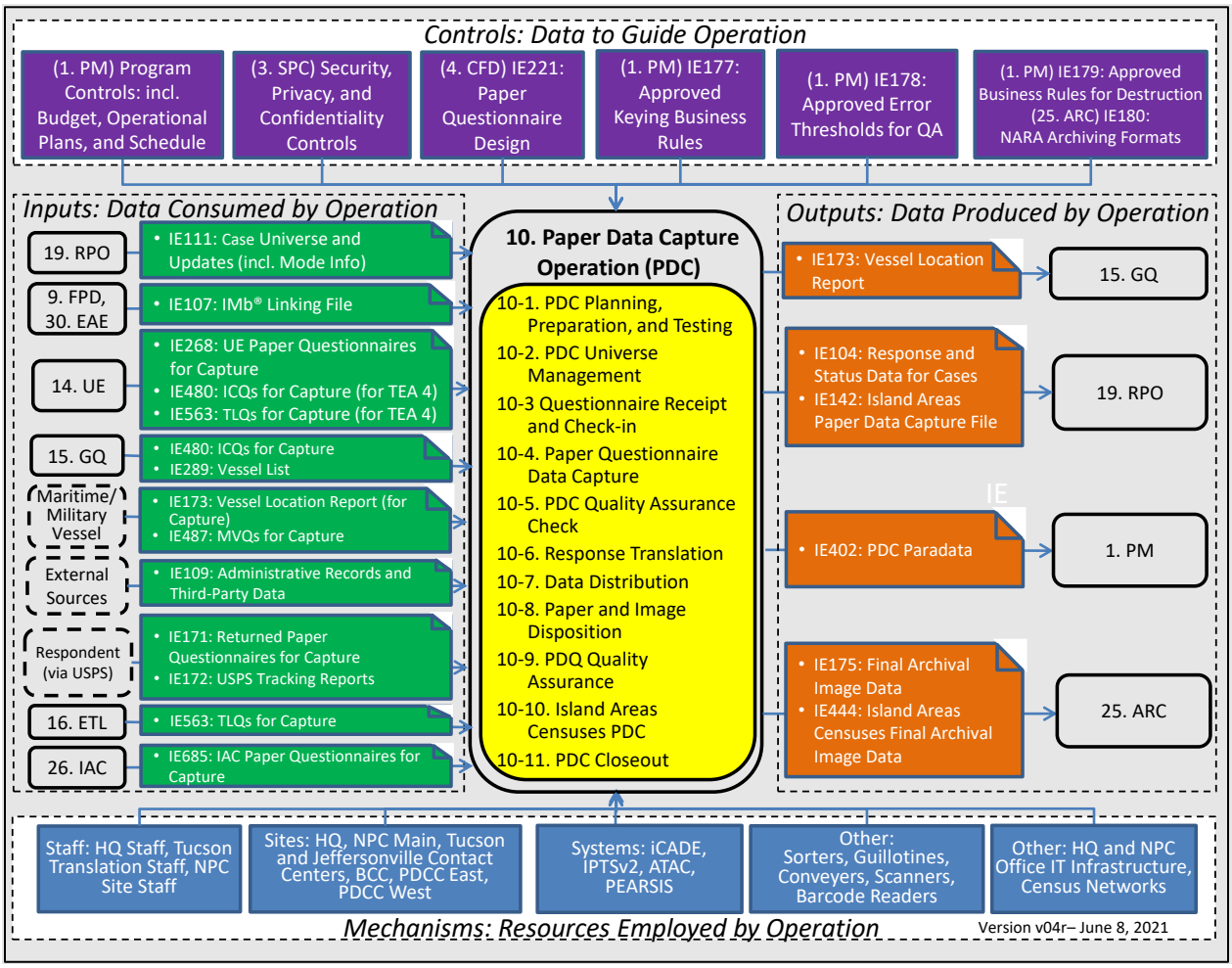

Figure 1: Paper Data Capture Operation (PDC) Context Diagram

PDC utilizes inputs from a variety of other operations internal and external to the Census Bureau. Before the actual paper data capture operation can begin, the initial mailing universe must be received by RPO for stateside and Puerto Rico (PR) questionnaires.

PDC dependencies include the following:

- **Workload universe** – For stateside and Puerto Rico (PR), both IMb® Confirm Service Postal Tracking System version 2 (IPTSv2) and Automated Tracking and Control System (ATAC) require the production universes to be received and loaded prior to processing.
- **Requirements & Specifications** – PDC solution providers need requirements defined to meet decennial and IAC needs.
- **USPS Status Data** – IPTSv2 is dependent on the USPS for tracking data.
- **IT Enterprise** – PDC is dependent on Decennial and HQ enterprise IT solutions.

- **Network** – PDC is dependent on the Telecommunications Office (TCO) for networking solutions and support.
- **Decennial Contracts and Execution Office (DCEO)** – PDC is dependent on the Field IT Deployment (FITd) contract for PDCC IT deployments and IT support; PDC is dependent on the Technical Integrator (TI) contract for IPTSv2, testing support, and Performance and Scalability (P&S) testing.
- **General Services Administration (GSA)** – PDC is dependent on the GSA for PDCC facility identification and acquisition.
- **NPC** – PDC is dependent on NPC for IT services, facility management, systems and program management support and the conduct of the operation.
- **Forms Printing and Distribution operation (FPD)** – PDC is dependent upon FPD for accurately printing and providing paper questionnaires in accordance with iCADE system requirements and specifications.
- **Unified Tracking System (UTS)** – PDC is dependent on UTS for reporting metrics.
- **Associate Director for Economic Programs (ADEP)** – PDC is dependent on the Economic Directorate to provide the iCADE paper data capture system.
- **POP** – PDC is dependent on Population Division (POP) to supply the capture and business rules for systems and operations.

For detailed descriptions of the Inputs, Controls, Outputs, and Mechanisms used by the PDC operation, see the sections that follow.

2.3.2.1 PDC Operational Inputs

Inputs are the data that are consumed by the operation. The inputs define the amount of operational work that needs to be performed.

Table 1 lists the inputs to the PDC operation.

Table 1: PDC Operational Inputs

Provider	Information Exchange	Description
19. Response Processing Operation (RPO)	IE111: Case Universe and Updates (including Mode Info)	The set of cases (i.e., living quarters) to be enumerated and the contact strategy (Internet First, Internet Choice, Update Leave) for each case. For PDC, this includes the entire self-response universe. An initial case universe is sent at the beginning of the PDC operation by RPO. Universe updates are reflected as work completes in all modes.
9. Forms Printing and Distribution operation (FPD) 30. Evaluation and Experiments operation (EAE)	IE107: IMb® Linking File	Results of the bulk mail contract arrangements with United States Postal Service (USPS) used to deliver mailing items. Each mailing item (package, letter, or postcard) has an Intelligent Mail barcode (IMb®) for postal tracing. IMb® Confirm Service Postal Tracking System version 2 (IPTSv2) needs to know the IMb® code and corresponding Census ID (or Document ID) that is associated with each item that is mailed out or returned by respondents to PDC. This information is used by PDC and RPO to update case status and to identify USPS Undeliverable As Addressed (UAA) mailing items.
14. Update Enumerate operation (UE)	IE268: UE Paper Questionnaires for Capture	Paper questionnaires completed as part of UE enumeration of housing units. These questionnaires have been prepared and are ready to be sent for processing at PDC.

Provider	Information Exchange	Description
14. Update Enumerate operation (UE)	IE480: ICQs for Capture (for TEA 4)	Individual Census Questionnaires (ICQ) collected from group quarters (GQs) in Type of Enumeration Area (TEA) 4. These ICQs have been prepared and are ready to be sent for processing at PDC. This is only for group quarters in TEA 4 - Remote Alaska.
14. Update Enumerate operation (UE)	IE563: TLQs for Capture (for TEA 4)	The Transitory Location Questionnaires (TLQs) are reviewed and logged for format and completeness (processed in the field) prior to being sent to PDC for scanning and data capture. For UE, this includes paper questionnaires completed as part of UE enumeration of transitory locations for TEA 4 Remote Alaska.
15. Group Quarters operation (GQ)	IE480: ICQs for Capture	Individual Census Questionnaires (ICQ) collected from GQs. These ICQs have been prepared and are ready to be sent for processing at PDC.
15. Group Quarters operation (GQ)	IE289: Vessel List	List of maritime/military vessels in operation that are applicable to the 2020 Census.
Maritime/Military Vessel	IE173: Vessel Location Report (for Capture)	A report that provides geographic location to be assigned to the 2020 Census results collected from a maritime/military vessel. This information is shipped with the Military Vessel Questionnaires (MVQs).
Maritime/Military Vessel	IE487: MVQs for Capture	MVQs that have been included in the mailout/mailback from the maritime or military vessel, sent directly to PDCC East for data capture.

Provider	Information Exchange	Description
External Sources	IE109: Administrative Records and Third-Party Data	Data from administrative records from other government sources or third-party data from commercial sources that are used by the Administrative Records modeling function to determine the occupied status of a living quarter (Occupied, Vacant, Delete, or Undetermined) and to determine the best time of day to contact the household to improve the likelihood of a successful contact attempt. PDC uses administrative records to assist in calibrating correctness of Optical Character Recognition (OCR).
Respondent (via United States Postal Service [USPS])	IE171: Returned Paper Questionnaires for Capture (from TEA 1 and TEA 6)	Paper questionnaires that were mailed back to the Census Bureau by respondents. Respondents in TEA 1 (self-response) and TEA 6 (Update Leave) can fill out questionnaires they receive and mail them via USPS to the Census Bureau for capture by the PDC operation.
Respondent (via USPS)	IE172: USPS Tracking Reports	Data provided by USPS related to mail delivery and undeliverable status of paper questionnaires. Delivery status data include the number of paper questionnaires en route to the Census Bureau and expected arrival dates.
16. Enumeration at Transitory Locations operation (ETL)	IE563: TLQs for Capture	The Transitory Location Questionnaires (TLQs) reviewed and logged for format and completeness (processed in the field) prior to being sent to PDC for scanning and data capture.

Provider	Information Exchange	Description
26. Island Areas Censuses operation (IAC)	IE685: IAC Paper Questionnaires for Capture	Paper questionnaires completed as part of IAC enumeration of housing units, group quarters, and transitory locations in the Island Areas. NPC staff ensure received questionnaires are ready to be sent for processing by PDC. IAC questionnaires mostly resemble the ACS questionnaire, so the ACS data capture system is used. IAC PDC is conducted at NPC Main Campus.

2.3.2.2 PDC Operational Controls

Controls are the data that guide the behavior of the operation. They are not consumed by the operation, but rather they provide guidance, models, limits, criteria, cutoff dates, or other information that controls the way in which the operational work is performed.

Table 2 lists the controls for the PDC operation.

Table 2: PDC Operational Controls

Provider	Information Exchange	Description
1. Program Management operation (PM)	Program Controls	Program Control information including: • Budget. • Operational Plans and Schedule.

Provider	Information Exchange	Description
3. Security, Privacy, and Confidentiality operation (SPC)	Security, Privacy, and Confidentiality Controls	Laws, policies, regulations, and guidelines related to physical security, IT security, data security, and privacy and confidentiality impacts, analyses, and processes. These include but are not limited to Title 13, Title 26, and other laws and policies related to protection of personally identifiable information.
4. Content and Forms Design operation (CFD)	IE221: Paper Questionnaire Design	Design of the paper questionnaires. Includes templates for each questionnaire form type to be used for setting up the scanners and software for Optical Mark Recognition (OMR), Optical Character Recognition (OCR), Key From Image (KFI), and Barcode Recognition (BCR).
1. Program Management operation (PM)	IE177: Approved Keying Business Rules	Approved business rules for keying specific fields and write-in responses while capturing data from the paper questionnaires. There are also capture rules for automated and key entry.
1. Program Management operation (PM)	IE178: Approved Error Thresholds for Quality Assurance (QA)	Approved percentage of allowed errors detected during quality review of a batch of sample fields. If the percentage of errors exceeds the threshold, the batch fails quality assurance and the remainder of the fields in the batch are rekeyed.
1. Program Management operation (PM)	IE179: Approved Business Rules for Destruction	Approved business rules that dictate when and how paper questionnaires can be destroyed after imaging.

Provider	Information Exchange	Description
25. Archiving operation (ARC)	IE180: National Archives and Records Administration (NARA) Archiving Formats	Acceptable archiving formats to assist in determining how to prepare response records for archiving based on NARA and Census Bureau requirements.

2.3.2.3 PDC Operational Outputs

Outputs are the data produced by the operation. The outputs constitute the results of operational work that has been performed. Outputs produced may be used as inputs or controls to other operations.

Table 3 lists the outputs from the PDC operation.

Table 3: PDC Operational Outputs

Consumer	Information Exchange	Description
15. Group Quarters operation (GQ)	IE173: Vessel Location Report	A report that provides geographic location to be assigned to the 2020 Census results collected from a maritime/military vessel. This information is shipped with the Military Vessel Questionnaires (MVQs).
19. Response Processing Operation (RPO)	IE104: Response and Status Data for Cases	Response data and associated status information that result from enumeration of cases in the 2020 Census Enumeration Case Universe. For PDC, the response data are captured from paper questionnaires that have been scanned and imaged. Examples of PDC status data include forms checked in through iCADE and Data Delivered.

Consumer	Information Exchange	Description
19. Response Processing Operation (RPO)	IE142: Island Areas Paper Data Capture File	Response and status information captured by PDC from paper questionnaires employed during the IAC data collection operations work.
1. Program Management operation (PM)	IE402: PDC Paradata	Status and progress data related to the PDC activities. May include counts on checkout or where a response is in the system.
25. Archiving operation (ARC)	IE175: Final Archival Image Data	Captured data from scanned paper images to be indexed and archived based on NARA requirements, including images for completed questionnaires for housing units and for other unit types such as group quarters (GQs) and transitory locations.
25. Archiving operation (ARC)	IE444: Island Areas Censuses Final Archival Image Data	Captured data from scanned paper images for Island Areas Censuses to be indexed and archived based on NARA requirements.

2.3.2.4 PDC Operational Mechanisms

Mechanisms are the resources (people, places, and things) that are used to perform the operational processes. They include staff resources, infrastructure sites, systems, and other technology infrastructure.

Staff Resources

Table 4 identifies the staff resources employed for the PDC operation.

Table 4: Staff Resources Used Within PDC Operational Activities

Staff Resources	Description/Role
Headquarters (HQ) Staff	All staff members located at Census Bureau HQ who support this operation. This includes integrated Computer Assisted Data Entry (iCADE) system staff at HQ who coordinate systems at the paper data capture centers as well as Decennial staff who support and report status of PDC activities. It may also include staff located at the Census Bureau's Bowie Computer Center (BCC).
Tucson Translation Staff	Staff who translate and key selected response fields that are not in English or Spanish for paper data capture operations.
NPC Site Staff	NPC staff who manage and operate the two PDCC sites and main campus. These include the ATAC, iCADE, clerks and laborers, and Operations Management staff, as well as people overseeing the receipt, scanning preparation, imaging, KFI processing, and paper questionnaire destruction. This also includes staff responsible for quality assurance and support services activities associated with the PDC operation.

Infrastructure Sites

Table 5 identifies the infrastructure sites employed for the PDC operation.

Table 5: Infrastructure Sites for PDC Operational Activities

Infrastructure Site	Description/Role
Headquarters (HQ)	HQ for office work. This location is in Suitland, Maryland.

Infrastructure Site	Description/Role
National Processing Center (NPC) Main Campus	The NPC Main Campus site is in Jeffersonville, Indiana, and is responsible for the following example activities: • Receiving completed paper data collection materials of the 2020 Island Areas Censuses (IAC). • Checking-in, sorting, batching completed IAC paper data collection materials from the 2020 Census. • Preparing IAC questionnaires for data capture. • Scanning and processing IAC questionnaires using integrated Computer Assisted Data Entry (iCADE) for surveys.
Tucson and Jeffersonville Contact Centers	Tucson, AZ, and Jeffersonville, IN, are the locations for these centers. The bilingual staff at the Tucson National Processing Center (NPC) Contact Center translate detected response characters in the Internet Self-Response (ISR) and Paper Data Capture response mode (iCADE). For the 2020 Census, response translation is limited to the supported non-English and non-Spanish languages and only for the Hispanic origin and race write-in fields.
Bowie Computer Center (BCC)	Census Bureau's BCC is in Bowie, Maryland. The BCC is responsible for the Backup/Recovery operations for the 2020 PDC operation. There are two backup copies of the PDC data. The first backup copy is at each PDCC site (Indiana and Arizona); the second backup copy for each site resides at the BCC.
Paper Data Capture Center (PDCC) East	A site at which paper data capture processing is performed. This site is in the Jeffersonville, Indiana area.
PDCC West	A site at which paper data capture processing is performed. This site is in the Phoenix, Arizona area.

Systems and Other Technology Infrastructure

Table 6 identifies the systems employed for the PDC operation.

Table 6: Systems Used Within PDC Operational Activities

System	Description
integrated Computer Assisted Data Entry (iCADE)	A data capture solution for paper-based data collection operations. For PDC, iCADE receives paper questionnaires and produces document images and form control information, provides OCR, OMR, and other image processing functions, and uses these to extract respondent data as well as to control KFI and QA.
IMb® Confirm Service Postal Tracking System version 2 (IPTSv2)	IPTSv2 is a Census Bureau system used to ingest data from the USPS Informed Visibility service. The system generates records for each mail piece when it is processed through an automated sort. These events can be correlated to an expected delivery date of outbound mail and provide information about business reply mail making its way to a data capture facility and can be used to inform PDC of potential workload spikes.
Automated Tracking and Control System (ATAC)	The ATAC system serves as the National Processing Center's operational control system for tracking and reporting activities involving survey materials, including those for the 2020 Census. It is used to inform NPC processing operations of the status and progress of current and ongoing activities, including: • Checking in received questionnaires. • Tracking clerical review. • Tracking material storage information. • Tracking shipping information of materials to Census PDC processing locations. • Reporting out on all aspects of information stored in the database. • Ensuring that the batches of paper questionnaires identified by iCADE as ready for destruction have in fact been delivered for destruction.

Production Environment for Administrative Records Staging, Integration and Storage (PEARSIS)	PEARSIS provides a repository of federal administrative records and third-party data.

Other technology infrastructure employed for the PDC operation includes:

- Sorters.
- Guillotines.
- Conveyers.
- Scanners.
- Barcode readers.
- HQ and NPC office information technology (IT) infrastructure.
- Census networks.

2.4 PDC Data Flow and Operational Influences

Figure 2 is an Integrated Operations Diagram (IOD), which describes the design concepts for the response data collection operations for the 2020 Census (stateside and Puerto Rico). This diagram assumes that the frame has been developed and address canvassing operations are complete. The diagram shows the Response Processing Operation (RPO) as the hub of data collection and RPO's interactions with all the other 2020 Census operations that have a role in data collection. The discussion below walks the reader through the diagram, using the circled numbers to help the reader follow the flow.

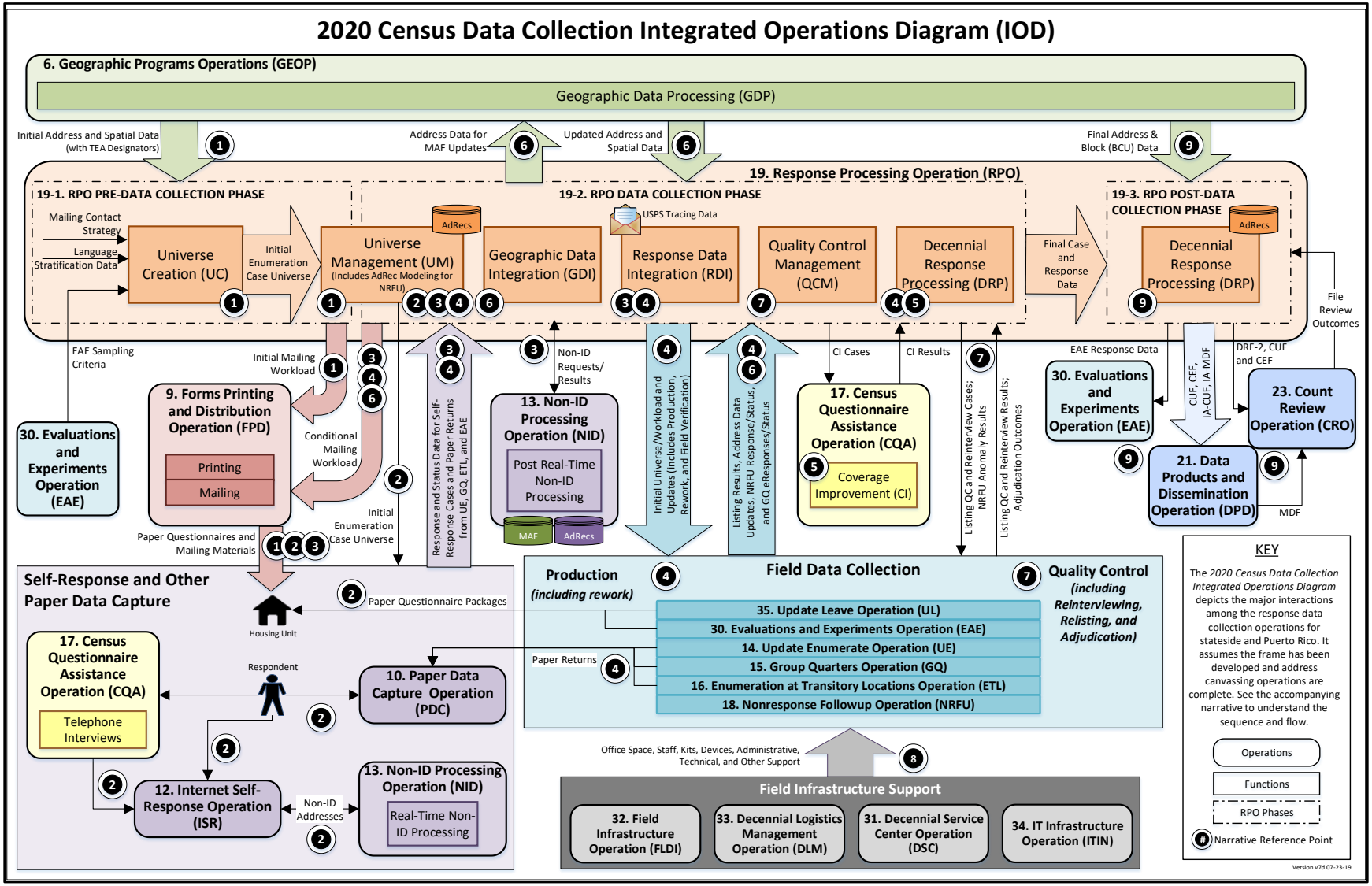

Figure 2: 2020 Census Data Collection - Integrated Operations Diagram (IOD)

Pre-Data Collection

Before the start of data collection, the Geographic Data Processing (GDP) component of the Geographic Programs operation (GEOP) sends initial Address and Spatial Data, including the Type of Enumeration Area (TEA) designations, to RPO so it can create the Initial Enumeration Case Universe. RPO also receives the mailing contact strategy (i.e., strategy for self-response stratification) so it can identify which housing units receive which kinds of mailings, language stratification information so it knows which language to use, and experimentation stratification data so it knows which housing units are to be included in what types of experiments. The creation of the Initial Enumeration Case Universe and application of the stratification data are done as part of the RPO Universe Creation function.

Based on the stratifications, the RPO Universe Management function creates the initial mailing workload and sends it to the Forms Printing and Distribution operation (FPD), which prints and then mails the appropriate materials to housing units for the Self-Response (SR) TEA. The first two of the five potential mailings for the SR TEA are sent unconditionally to all housing units in this TEA. These mailings are sent in English or English and Spanish based on the language stratification data and may include letters or—based on the self-response stratification—questionnaires.

During Data Collection

Once the RPO Universe Creation work is complete, the Initial Enumeration Case Universe is managed by the RPO Universe Management function, which tracks changes to the enumeration universe for future mailings and for the data collection operations.

People living in housing units are encouraged to self-respond through a partnership and communications campaign (not shown on this diagram), through mailings sent by FPD, and through paper questionnaires left at housing units as part of the Update Leave operation (UL).

To make it easy for people to respond and to reduce the paper workload, the Census Bureau is using an *Internet First* strategy for most housing units. Respondents can go to the internet and enter their response using the internet instrument as part of the Internet Self-Response operation (ISR). The internet option offers additional flexibility and allows people to respond in multiple languages. If a respondent calls the Census Questionnaire Assistance operation (CQA), a customer service representative may offer to collect the respondent's information by telephone. The information collected from these telephone interviews is entered by a customer

service representative using an ISR instrument similar to the public-facing instrument used by respondents.

Respondents can also mail paper questionnaire forms. These forms are received by the Paper Data Capture operation (PDC), which uses scanning and imaging technology to capture the information from these forms.

ISR receives the Initial Enumeration Case Universe from the RPO Universe Management function and uses the Enumeration Case Universe to link responses provided through the internet instrument to the appropriate case. If respondents do not have their unique Census ID available, they are still able to complete the census questionnaire as a Non-ID response using the ISR instrument. The Non-ID Processing operation (NID) first attempts to match the address entered by the respondent or customer service representative to a known census address in real-time. For addresses that do not match, the response is still collected and is subject to later Non-ID Processing.

Response and status data collected through the various self-response data collection operations are sent (in digital format) to RPO's Response Data Integration function. Any responses collected through PDC or ISR that are submitted in languages other than English or Spanish are translated by staff at the Tucson call center on behalf of these operations before being sent to RPO. RPO's Universe Management function uses the response status data to determine the appropriate actions for the case.

During the self-response data collection time-period, reminder mailings are sent to housing units in the SR TEA. The first reminder is sent to all housing units in the SR TEA using the initial mailing workload as discussed above. Subsequent reminders are conditional and are only sent to those housing units that have not yet responded. The RPO Universe Management function sends a Conditional Mailing Workload to the FPD operation for these nonresponding units. FPD also receives from the RPO Universe Management function a list of mailable housing units in the UL TEA and mails two reminders to these housing units.

Any remaining Non-ID cases are sent by RPO to NID for post real-time Non-ID processing, which attempts to match addresses provided by respondents to known addresses in the Master Address File (MAF) using automated and clerical procedures. As needed, administrative records (AdRecs) are used to supplement the matching process. Most of these Non-ID cases are from internet responses that could not be matched during real-time Non-ID processing. The results of post real-time Non-ID matching are sent back to RPO. Based on predefined business rules,

some of the responses that are not able to be matched through NID are sent to the field for verification as part of the Nonresponse Followup operation (NRFU).

The discussion above covers self-responses for people living in housing units. Special operations also exist to collect data from people living in other types of living quarters or for whom self-response is not a viable option:

- The Group Quarters operation (GQ) enumerates people living in group quarters (e.g., dormitories, correctional facilities, and nursing/skilled-nursing facilities) as well as people experiencing homelessness and receiving services at service-based locations such as soup kitchens. GQ also enumerates people living on maritime and military vessels and living in group quarters on military bases and other military installations using specialized procedures.
- The Enumeration at Transitory Locations operation (ETL) enumerates people who are living in special locations—such as recreational vehicle parks, campgrounds, racetracks, circuses, carnivals, marinas, hotels, and motels—and who do not have a Usual Home Elsewhere.
- The Update Enumerate operation (UE) lists and enumerates housing units in areas that pose unique challenges to the standard self-response data collection operations. These housing units are in the UE and Remote Alaska TEAs, which cover remote areas of the country and other small, selected areas.

NRFU is another special operation whose primary purposes are to determine the housing unit status of addresses in the SR and UL TEAs for which a self-response was not received and to enumerate those that are believed to be occupied. As mentioned in number 3 above, NRFU also performs a field verification activity to verify selected addresses for Non-ID self-responses that could not be matched to known addresses through NID.

Based on the universe case type (derived from TEA and living quarter type), RPO sends the Initial Enumeration Case Universe/Workload to the GQ, ETL, and UE operations. GQ uses this universe to perform an advance contact activity to collect general information and determine the preferred method of enumeration. ETL also performs an advance contact activity to schedule appointments for enumerating its universe of cases.

NRFU does not require advance contact activities. For NRFU, the RPO Universe Management function creates an Initial Case Universe/Workload comprising all housing units in the SR and UL TEAs for which a self-response was not received. The NRFU contact strategy is dependent on

the results of an AdRec modeling activity. Four possible status outcomes result from this modeling for a given address:

- AdRec Vacant: No one lives there.
- AdRec Delete: There is no housing unit at that address.
- AdRec Occupied: There is a high probability that someone lives there and that the Census Bureau has high-quality data about that housing unit.
- AdRec No Determination: Administrative data are not sufficient to help determine the housing unit status, and a full contact strategy is required.

An initial attempt to enumerate is made for all addresses in the NRFU workload. NRFU sends to RPO information regarding the success of this and any subsequent enumeration attempts as part of the response status data.

AdRec Vacant and AdRec Delete housing units also receive an additional mailing from FPD. The RPO Universe Management function provides this AdRec mailing workload to FPD (as another type of Conditional Mailing).

RPO removes from the follow-up workload any AdRec Occupied cases that cannot be successfully resolved during the first attempt and triggers one final mailing (from FPD) to these addresses to encourage these households to self-respond. RPO also removes AdRec Vacant and AdRec Delete cases that cannot be successfully resolved during this first attempt, provided that those cases do not appear occupied and that information from the United States Postal Service (USPS) about the additional mailing indicates that the unit is either vacant or nonexistent. For all other cases, NRFU continues to attempt to enumerate the housing unit until the attempt is successful or the maximum number of attempts has been reached. Additional attempts are made for selected units during the NRFU Closeout phase based on a reassessment of the AdRec modeling results using a relaxed, lower quality threshold.

Self-responses can continue to arrive at any time during NRFU. Accordingly, RPO flags housing units in the follow-up workload for which RPO has received a self-response or tracing information from the USPS that indicates that a return is on its way to one of the paper data capture facilities. NRFU is notified about these flagged households as soon as the information is available so that it can remove the housing units from the daily workload, if possible. Any self-responses that are flagged but later found by RPO to have insufficient enumeration data are added back to the NRFU workload for continued enumeration attempts. The RPO Universe Management function tracks this information and uses it to determine what to include in the next day's follow-up workload. Housing units that have been successfully enumerated are not included in subsequent follow-up workloads.

For NRFU, field data are collected by electronic devices. The electronic data are sent to the RPO Response Data Integration function, which subsequently provides this information to the Decennial Response Processing function for further processing. Paper questionnaires are used to enumerate at living quarters during UE and at housing units during ETL. These paper questionnaires are checked-in at area census offices (ACOs) and then sent to the paper data capture facilities, where they are scanned and imaged by PDC. PDC sends the captured data and case status information to RPO in digital format.

GQ is primarily a paper operation. GQ responses collected on paper questionnaires are checked-in at the ACOs and sent to PDC for scanning and imaging before being transmitted to RPO. Case status updates are sent to RPO as part of the check-in process. GQ data provided in electronic files (eResponses) require additional processing to prepare the data before they are electronically transmitted to RPO. GQ data collected on paper rosters are entered by ACO clerks into the same file format that is used for eResponses. These response data are then sent electronically to RPO.

As part of the Evaluations and Experiments operation (EAE), the Census Bureau may test different questionnaire content and data collection methodologies during the 2020 Census to help evaluate content and modes for the 2020 Census and inform design changes for the 2030 Census. Addresses that are selected to be part of these experiments are identified in the initial universe (see number 1 above). For those addresses, the EAE operation sends households questionnaires with various experimental items, packaging, etc. to assess the impact made on the response or lack thereof. Respondents receiving EAE questionnaires and notices will respond via ISR, CQA, PDC, or NRFU.

5 RPO's Decennial Response Processing function performs quality assurance activities as well as coding and other preparation steps on incoming self-response data.

The RPO Universe Management function also supports a Coverage Improvement (CI) activity, the goal of which is to ensure a high-quality census by conducting telephone follow-up for households where there could be coverage issues on submitted responses. CI is a follow-up activity and is therefore considered a component of NRFU, however, the CI telephone interviews are performed by CQA. CQA receives from the RPO Universe Management function a set of cases with potential coverage issues and provides the results of these cases back to RPO's Response Data Integration function.

As noted above, universe and address updates occur during field operations. Census Bureau field staff may uncover changes to addresses as they perform their daily assignments in any field operation. For example, a UL or UE lister may add an address or find an error in the address or geographic data based on the listing activities, or a NRFU enumerator or a UL lister may go to an address and find an additional unit such as a garage apartment located on the premises. All listing results and other address changes are sent to the RPO Geographic Data Integration function, which passes the information on to the GDP function in GEOP.

Changes to the address list may also come from other sources such as appeals from the Local Update of Census Addresses operation (LUCA), the review of addresses performed by the Count Review operation (CRO), and updated files from the postal service. The GDP function within GEOP updates the address data and sends these to RPO's Universe Management Function, which provides these cases to the appropriate operation. Depending on the timing, living quarter type, and TEA designation, RPO may initiate one or more mailings to these new addresses through FPD to encourage self-response.

All field operations (GQ, UL, ETL, UE, and NRFU) include quality control (QC) functions.

For GQ, the field operational control system creates a sample of the field enumeration cases for QC by ACO staff, who conduct telephone reinterviews for this sample set of cases to confirm that a GQ enumerator visited the site and that the total population count is correct.

For UL, the RPO Quality Control Management function selects a sample set of basic collection units (BCUs) for relisting. The QC Listing Results are sent back to the RPO Quality Control Management function for further processing. RPO does not send any changes resulting from UL listing to GEOP until the lister has passed the QC check. Further, if the QC activities result in a hard fail, BCUs already worked may require relisting. RPO includes this rework in subsequent UL production workloads.

QC methods for ETL and UE are performed primarily in the field, tailored to meet the circumstances of these unique paper-based operations.

NRFU includes multiple methods for ensuring high-quality data collection. Several of these are integrated into the staff management activities. In addition, samples of field follow-up cases are selected for reinterview (RI), a process whereby the response data are collected again and compared with the original collected data. The RPO Quality Control Management function creates the RI workload and sends it to the NRFU operation. The RI cases are handled by NRFU field staff. The RI results are sent to the RPO Quality Control Management function, which

performs an automated comparison of the RI data against the original data. Anomalies are sent back to NRFU, where additional research is conducted to determine how these cases should be handled. The results of this review (adjudication outcomes) are sent back to the RPO Quality Control Management function. In some cases, the adjudication requires that prior cases performed by the enumerator at fault be reworked. RPO puts these cases back into the NRFU workload as appropriate.

NRFU, UE, UL, ETL, and parts of the GQ operation are performed in the field. Several operations provide the support for these field data collection activities. The Field Infrastructure operation (FLDI) recruits, hires, onboards, and trains the staff needed to conduct these operations and also operates the field offices during production. The Decennial Logistics Management operation (DLM) provides the space and logistics support (e.g., supplies and kits) for the offices and the field staff. The Decennial Service Center operation (DSC) provides technical support for field and field office staff. Finally, the IT Infrastructure operation (ITIN) provides the hardware and software used by the field staff and field offices.

Post-Data Collection

Once data collection is complete, additional processing occurs to prepare the counts for use in apportionment and the data used by the Data Products and Dissemination operation (DPD) to create data products for redistricting and other purposes. The RPO Decennial Response Processing function handles this post-data collection processing, which includes multiple activities:

- Supplementing response data with administrative records for those cases that had been identified as AdRec Occupied but for which a nonresponse followup attempt was unsuccessful and no subsequent self-response was received.
- Determining the final enumeration universe by reconciling or applying final address and block data from the GDP component of GEOP.
- Determining the returns of record for situations where multiple responses have been received for the same housing unit.
- Performing count and status imputations.
- Performing consistency editing and characteristic allocation supplemented with administrative records data.
- Applying tabulation geography.

- Performing disclosure avoidance. (Note: This is done by RPO for Island Areas Censuses data only. Disclosure avoidance for stateside/PR data is handled by DPD).

Similar processing occurs for responses from GQ. Responses collected through the EAE operation may require slightly different activities.

Through these processing activities, the RPO Decennial Response Processing function creates multiple files for stateside/PR response data, including the Decennial Response File (DRF), the Census Unedited File (CUF), and the Census Edited File (CEF). RPO also creates an Island Areas CUF (IA-CUF), Island Areas CEF (IA-CEF), and Island Areas Microdata Detail Files (IA-MDF) for the Island Areas Censuses (IAC) response data. Each of these files is reviewed within the Census Bureau before the data are sent to the next stage of processing. Some of these reviews are done as part of the CRO. The CUF, the CEF, the IA-CUF, and the IA-MDF are sent to DPD via the Census Data Lake (CDL). DPD uses these files as inputs for data products creation and also creates the stateside/PR MDFs using the CEF as input. RPO also sends data collected as part of EAE back to the EAE operation for further analysis.

3. Paper Data Capture Operation (PDC) Detailed Process Description

Figure 3 is a top-level Business Process Model (BPM) showing the Level 1 activity areas within the PDC operation. BPMs for the 2020 Census follow industry-standard Business Process Model and Notation (BPMN). An explanation of how to read the BPMN notations and a full-sized copy of all the BPMN diagrams for this operation are provided under separate cover.

This top-level BPM serves as the Context Model for the PDC operation. A BPMN Context Model displays the high-level activities within the operation and relationships between them, whereas the IDEF0 Context Diagram shown earlier depicts the boundaries of the operation or activity and the interfaces between the operation or activity and other operations and activities with which it is associated.

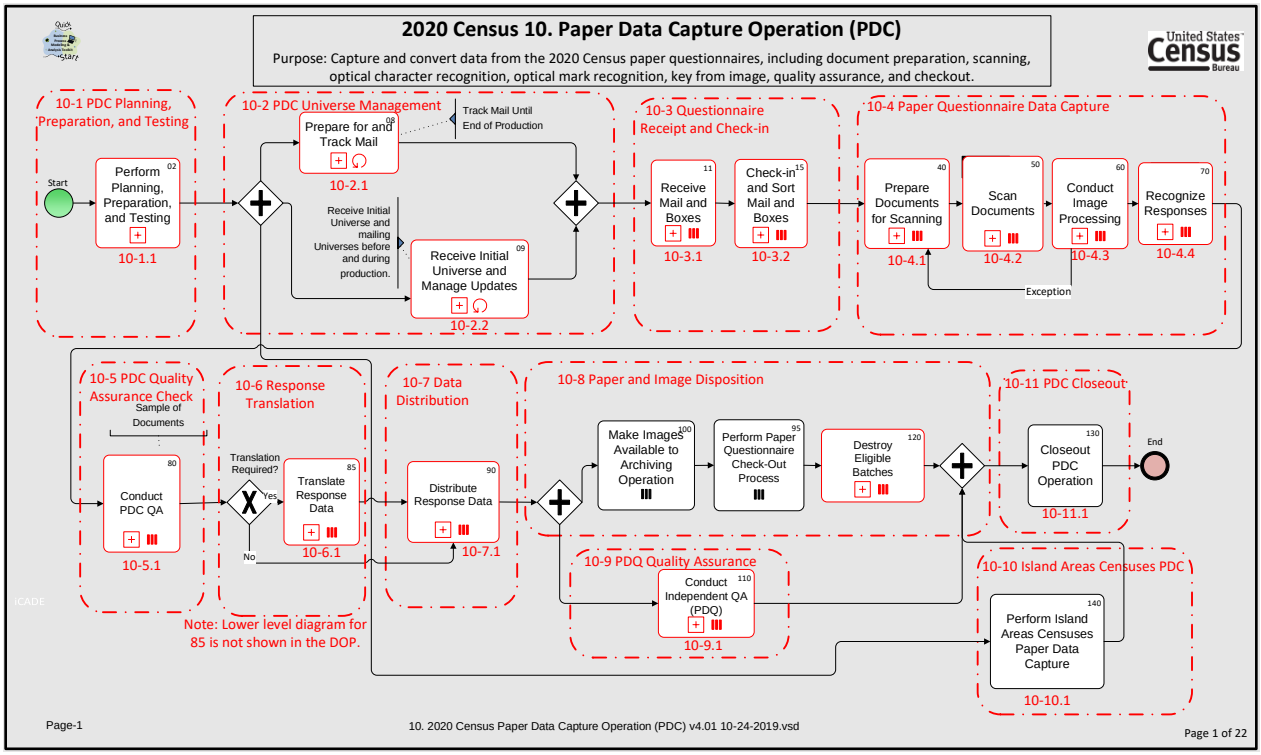

Figure 3: PDC Operation Context Model

The PDC operation is subdivided into the following activity areas:

- PDC Planning, Preparation, and Testing [PDC 10-1].
- PDC Universe Management [PDC 10-2].
- Questionnaire Receipt and Check-in [PDC 10-3].
- Paper Questionnaire Data Capture [PDC 10-4].

- PDC Quality Assurance Check [PDC 10-5].
- Response Translation [PDC 10-6].
- Data Distribution [PDC 10-7].
- Paper and Image Disposition [PDC 10-8].
- PDQ Quality Assurance [PDC 10-9].
- Island Areas Censuses PDC [PDC 10-10].
- PDC Closeout [PDC 10-11].

The business processes for each of these Level 1 activity areas are discussed along with their inputs and outputs in the following subsections.

3.1 PDC Planning, Preparation, and Testing [PDC 10-1]

Figure 4 shows the BPM for the PDC Planning, Preparation, and Testing [PDC 10-1] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

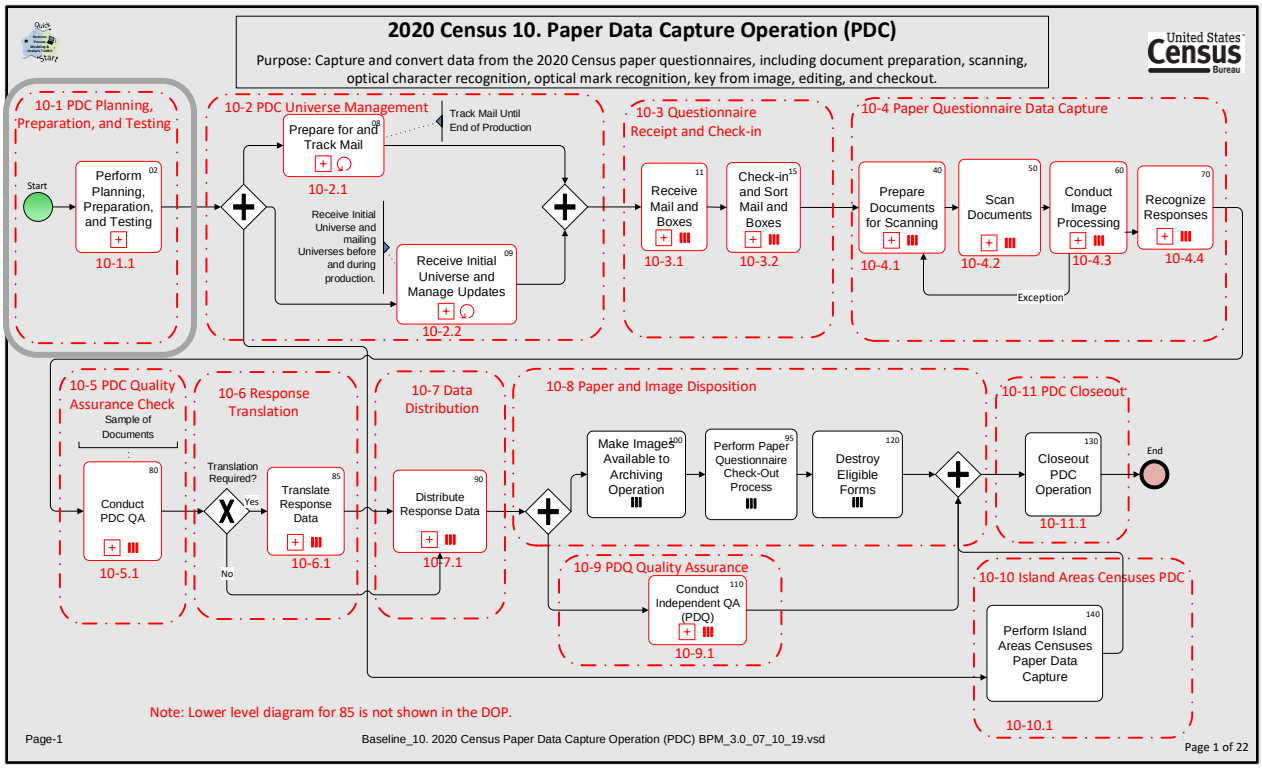

Figure 4: PDC Planning, Preparation, and Testing Constituent Activities [PDC 10-1]

The PDC Planning, Preparation, and Testing activity area consists of a single operational subactivity.

- PDC Planning, Preparation, and Testing [PDC 10-1].
    - Perform Planning, Preparation and Testing [PDC 10-1.1].

PDC assists each PDCC and NPC Main Campus with a series of tests that includes but is not limited to equipment testing, systems testing, facilities preparation, and staff training to ensure all operations run as planned when 2020 Census production begins. The next section describes the PDC Planning, Preparation, and Testing operational subactivity in detail.

3.1.1 Perform Planning, Preparation, and Testing [PDC 10-1.1]

A detailed view of the constituent activities that make up the "Perform Planning, Preparation, and Testing" operational subactivity is given in Figure 5 below.

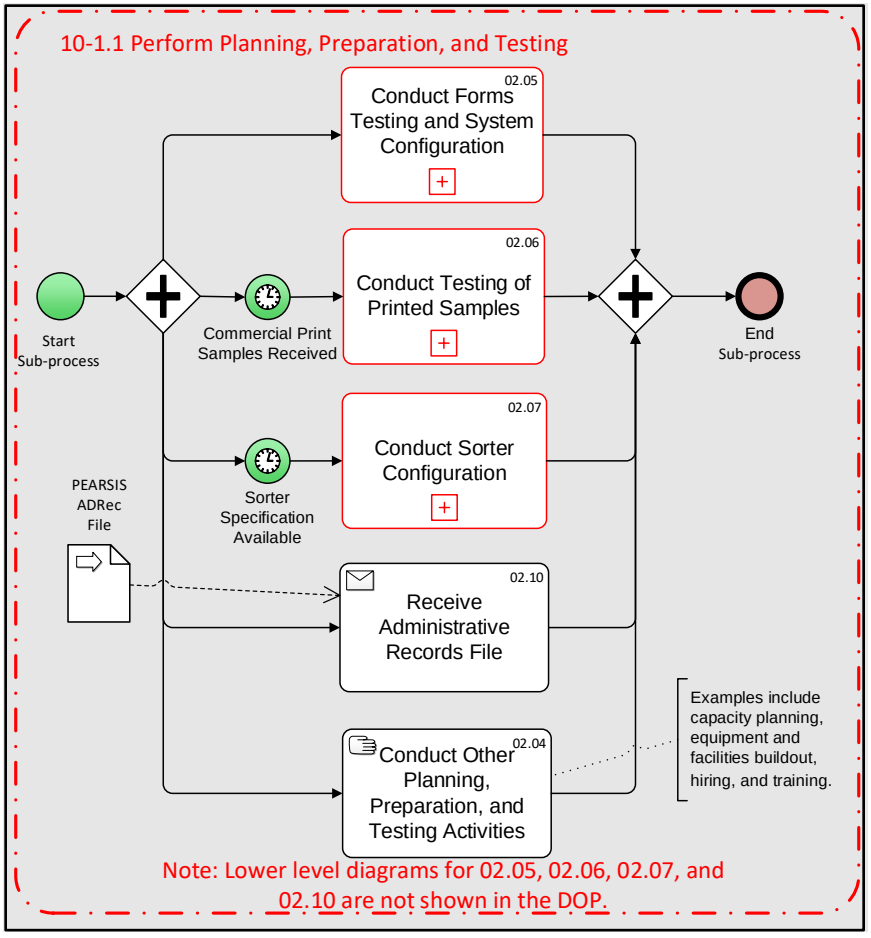

Figure 5: Perform Planning, Preparation, and Testing

PDC performs activities needed to plan and prepare for production. This includes testing activities, such as obtaining printed samples and questionnaires; system and sorter

configuration activities; capacity planning, equipment, and facilities buildout; and hiring and training of staff.

The PDCCs and NPC Main Campus plan to conduct preproduction testing to identify and resolve issues that could occur during production. These tests include Test Readiness Review (TRR), User Acceptance Test (UAT), Operational Test & Dry Run (OTDR), Performance & Scalability (P&S), Run for Record, Production Readiness Review (PRR), and Edit Deck test and review. Concerns and issues are documented along the way and affected systems and processes are retested to ensure accuracy.

3.2 PDC Universe Management [PDC 10-2]

Figure 6 below shows the BPM elements for the PDC Universe Management [PDC 10-2] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

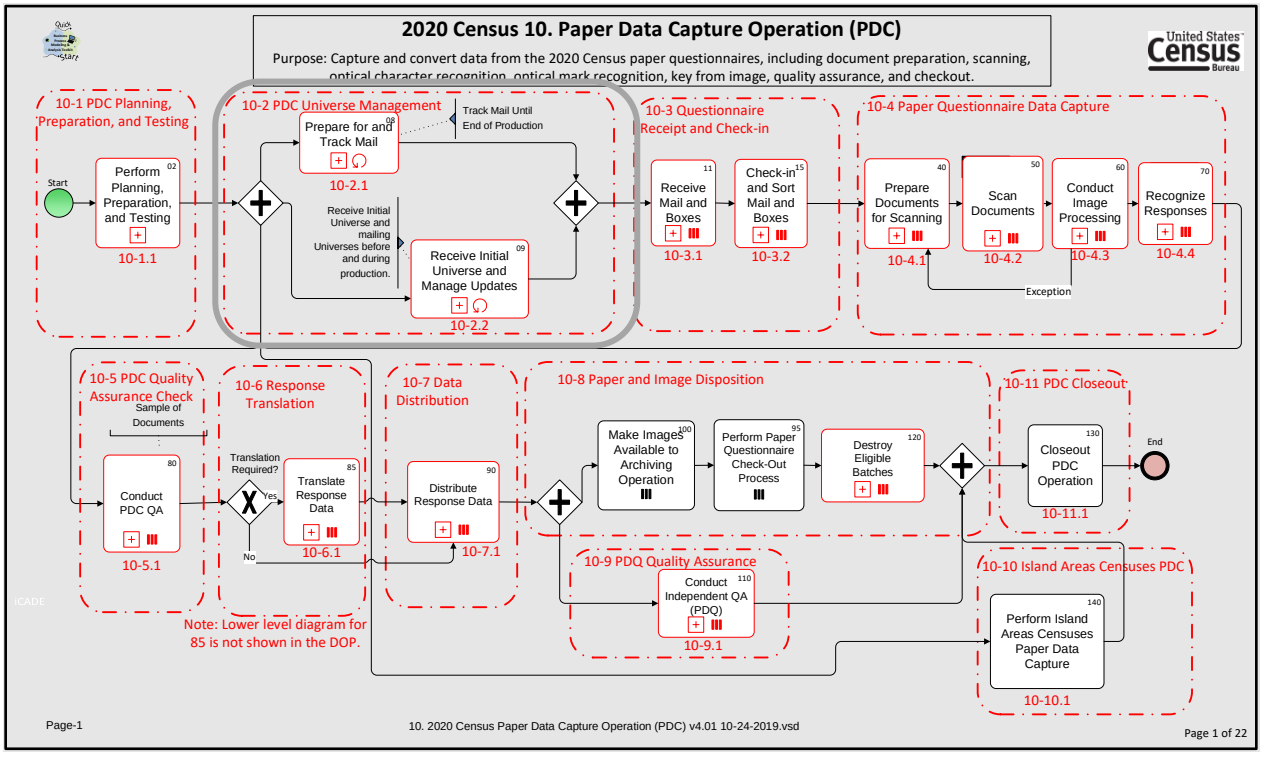

Figure 6: PDC Universe Management [PDC 10-2]

The PDC Universe Management activity area is subdivided into the following operational subactivities.

- Prepare for and Track Mail [PDC 10-2.1].
- Receive Initial Universe and Manage Updates [PDC 10-2.2].

One of the first and most important processes that kicks off the 2020 Census is the mailout and mailback operation. The IPTSv2 system tracks outgoing and incoming mail from the start of production until the end, utilizing the entire enumeration universe from RPO. The PDC operation, including the IPTSv2 system, receives the mail universe from RPO before questionnaire mailouts for all paper data capture operations. Throughout the production process, RPO provides updates to the universe as work completes in all modes to ensure each questionnaire is accounted for.

Subsequent sections describe the PDC Universe Management operational subactivities in detail.

3.2.1 Prepare for and Track Mail [PDC 10-2.1]

A detailed view of the constituent activities that make up the "Prepare for and Track Mail" operational subactivity is given in Figure 7 below.

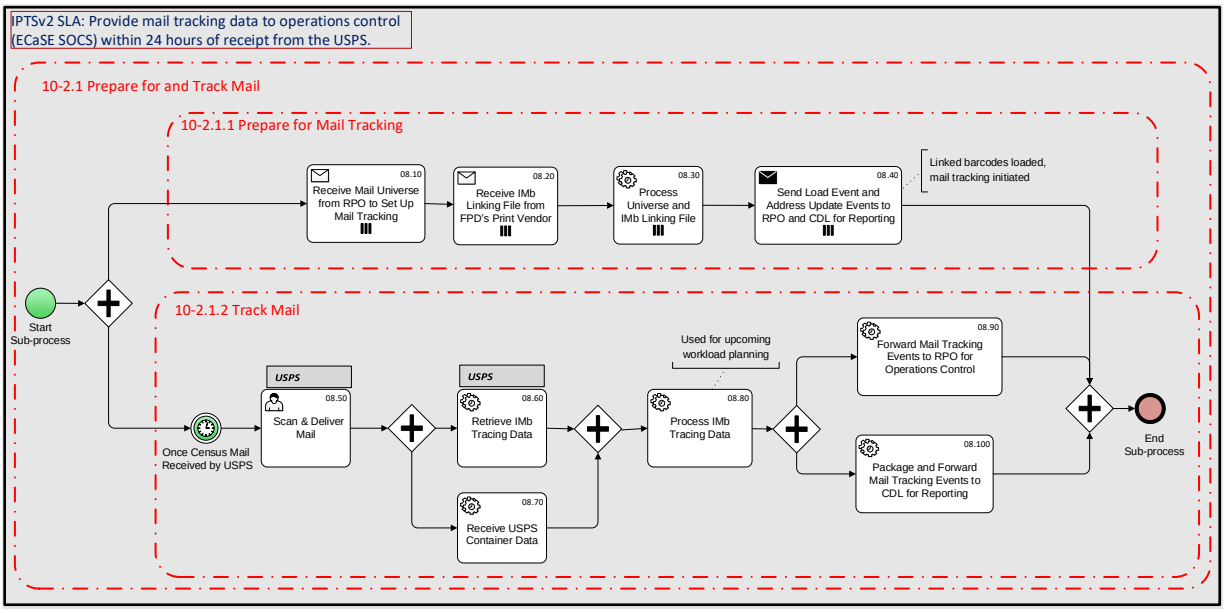

Figure 7: Prepare for and Track Mail

The Prepare for and Track Mail activity area is subdivided into the following operational subactivities:

- Prepare for and Track Mail [PDC 10-2.1].
  - Prepare for Mail Tracking [PDC 10-2.1.1].
  - Track Mail [PDC 10-2.1.2].

The subsequent sections describe the Prepare for and Track Mail operational subactivities in detail.

3.2.1.1 Prepare for Mail Tracking [PDC 10-2.1.1]

A detailed view of the constituent activities that make up the "Prepare for Mail Tracking" operational subactivity is given in Figure 8 below.

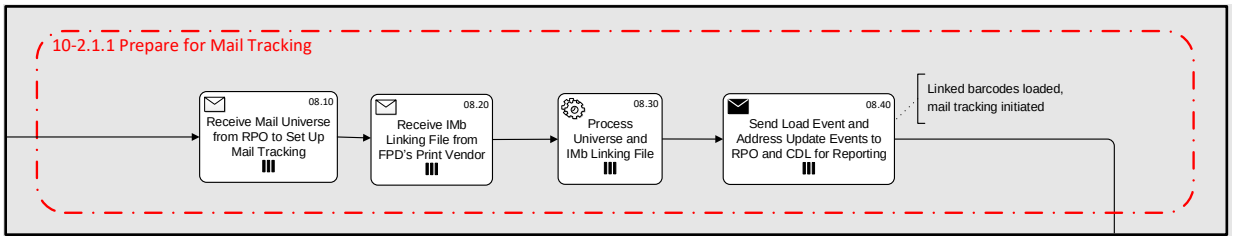

Figure 8: Prepare for Mail Tracking

The Prepare for Mail Tracking subactivity details the flow that supports the tracking of USPS mail both for outbound mailings to respondents and inbound returns with expected response data.

- First, universe files from RPO are loaded.
- Next, the print vendor creates mail-out packages and creates the IMb® linking file and PDC receives the file.
- When packages are mailed to respondents and when respondents mail back questionnaires, USPS tracking data become available and are delivered to IPTSv2 for processing.
- IPTSv2 sends results to the RPO and CDL for use.

3.2.1.2 Track Mail [PDC 10-2.1.2]

A detailed view of the constituent activities that make up the "Track Mail" operational subactivity is given in Figure 9 below.

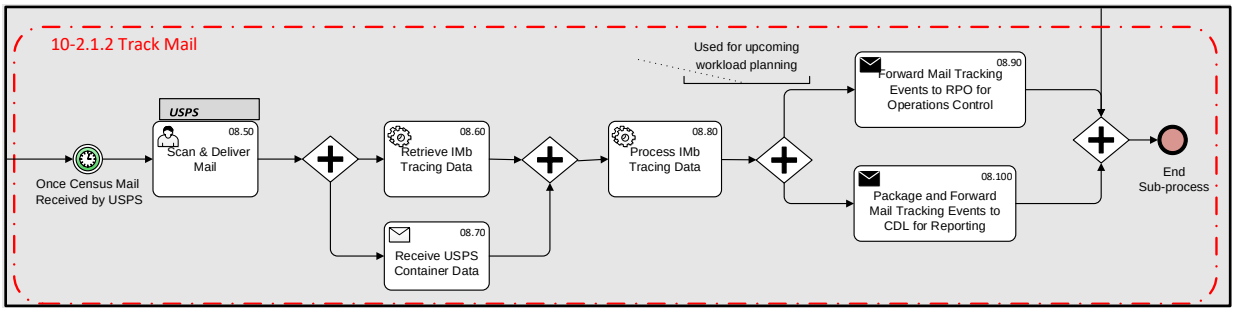

Figure 9: Track Mail

The Track Mail subactivity starts off with IPTSv2 in production processing USPS tracking data.

- The first step is for the USPS to scan mail and deliver data.
- Next step is for IPTSv2 to retrieve that data and process.
- IPTSv2 delivers data.

3.2.2 Receive Initial Universe and Manage Updates [PDC 10-2.2]

A detailed view of the constituent activities that make up the “Receive Initial Universe and Manage Updates” operational subactivity is given in Figure 10 below.

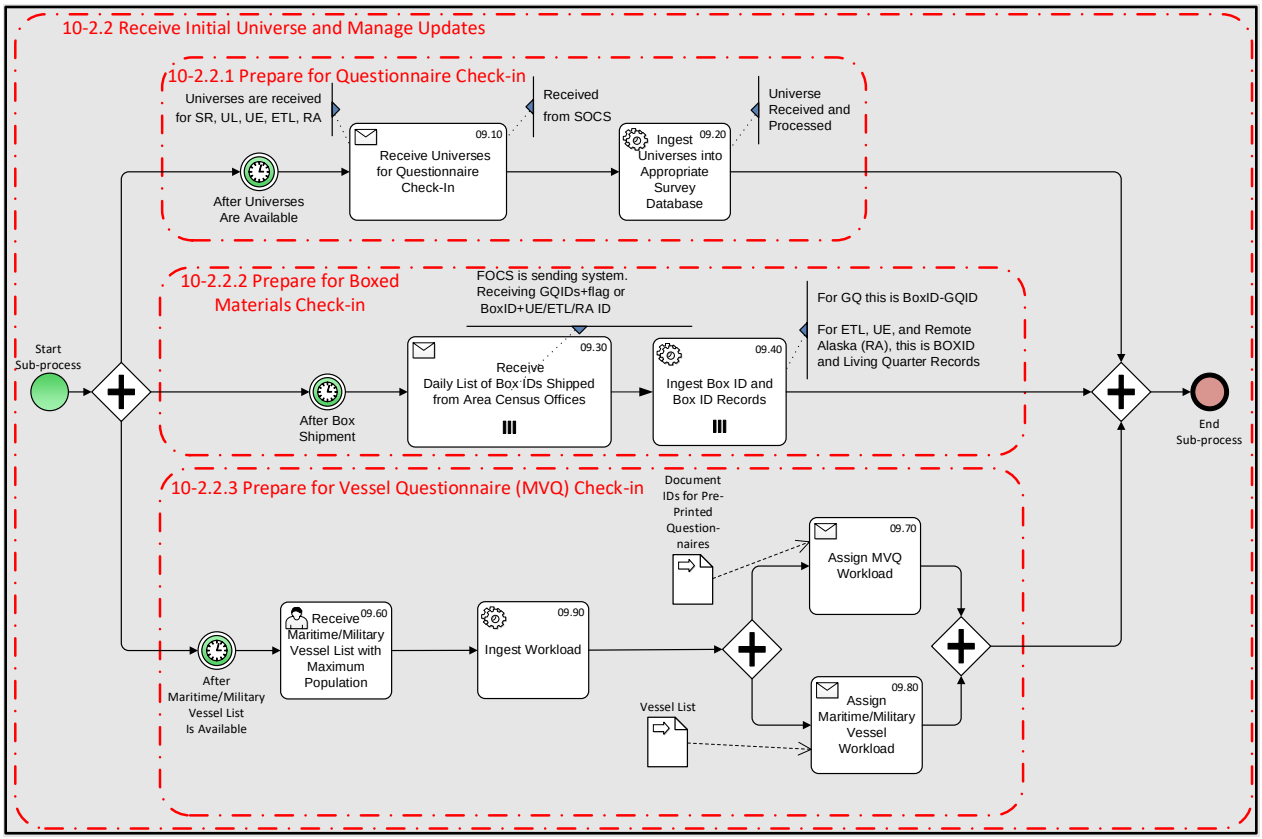

Figure 10: Receive Initial Universe and Manage Updates

The Receive Initial Universe and Manage Updates activity area is subdivided into the following operational subactivities:

- Receive Initial Universe and Manage Updates [PDC 10-2.2].
  - Prepare for Questionnaire Check-in [PDC 10-2.2.1].
  - Prepare for Boxed Materials Check-in [PDC 10-2.2.2].
  - Prepare for Vessel Questionnaire (MVQ) Check-in [PDC 10-2.2.3].

The PDC operation receives the entire universe of cases for which paper forms could be captured from RPO before the questionnaire mailouts and other paper enumeration operations.

The subsequent sections describe the Receive Initial Universe and Manage Updates subactivities in detail.

3.2.2.1 Prepare for Questionnaire Check-in [PDC 10-2.2.1]

A detailed view of the constituent activities that make up the "Prepare for Questionnaire Check-in" operational subactivity is given in Figure 11 below.

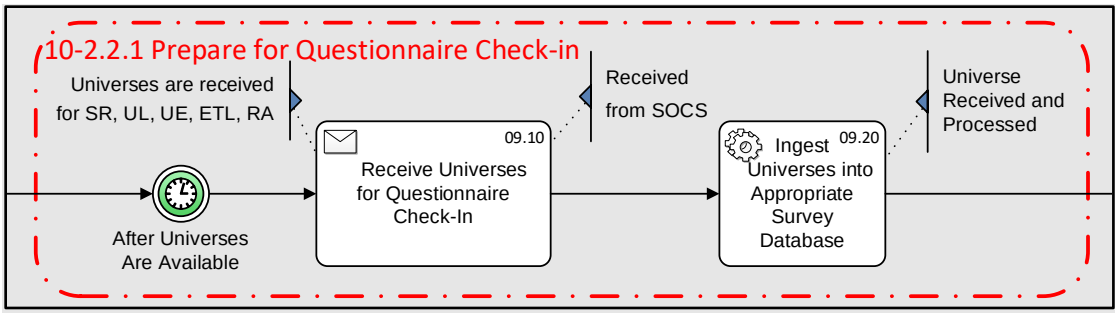

Figure 11: Prepare for Questionnaire Check-in

The PDC operation receives the initial universe of cases for which paper forms could be captured from RPO before the questionnaire mailouts and other paper enumeration operations. The universe of cases is continually updated throughout the operation through updates received from RPO.

Upon receiving the initial PDC Universe, the PDC operation creates a PDC workload that represents all the cases that could potentially require paper data capture. These inputs include universes for self-response, UL, ETL, and UE including Remote Alaska.

The subactivity for Ingest Universe into Appropriate Survey Databases represents the process of ATAC also loading the production universe. Each received production case must first be in the universe to be checked in or it is rejected.

3.2.2.2 Prepare for Boxed Materials Check-in [PDC 10-2.2.2]

A detailed view of the constituent activities that make up the "Prepare for Boxed Materials Check-in" operational subactivity is given in Figure 12 below.

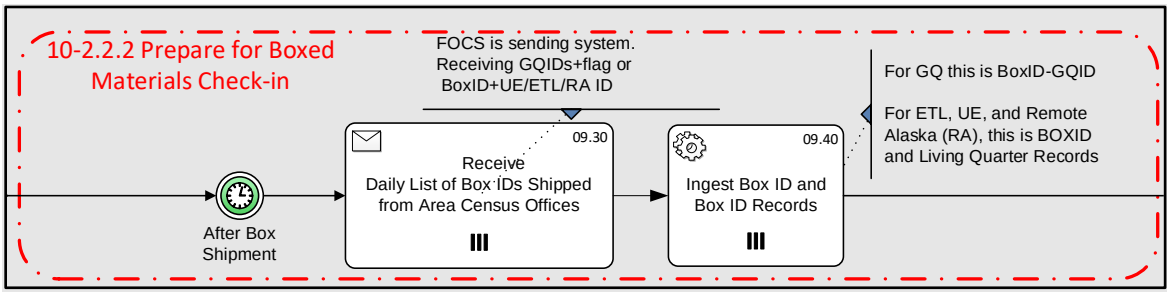

Figure 12: Prepare for Boxed Materials Check-in

The Prepare for Boxed Materials Check-in subactivity represents the process flow of ATAC receiving the box ID check-in universe from Field Operation Control System (FOCS). The activity begins after PDC receives the box shipment ID list from FOCS. The box ID must be in the ATAC universe for the box to be checked in once physically received at the PDCC. If the box cannot be checked in, all questionnaires inside the box are still processed with the box ID remaining in the "not received" status. Boxes are shipped from the GQ, ETL, and UE field activities including Remote Alaska.

3.2.2.3 Prepare for Vessel Questionnaire (MVQ) Check-in [PDC 10-2.2.3]

A detailed view of the constituent activities that make up the "Prepare for Vessel Questionnaire (MVQ) Check-in" operational subactivity is given in Figure 13 below.

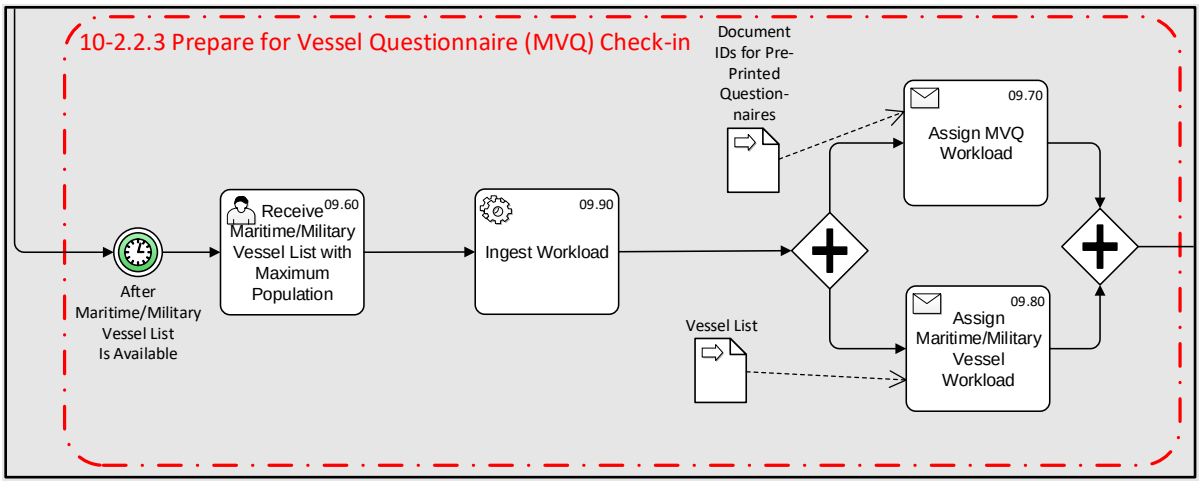

Figure 13: Prepare for Vessel Questionnaire (MVQ) Check-in

Prepare for Vessel Questionnaire Check-in represents the special path that MVQ questionnaire work takes. PDC receives a maritime/military vessel list from GQ that contains an estimate of the maximum associated population. PDC's system (ATAC) ingests the data for PDC to create

assignments for the vessels and the associated workload of expected MVQs. Inputs for the activity include vessel lists and document IDs for preprinted MVQs.

3.3 Questionnaire Receipt and Check-in [PDC 10-3]

Figure 14 below shows the BPM for the Questionnaire Receipt and Check-in [PDC 10-3] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

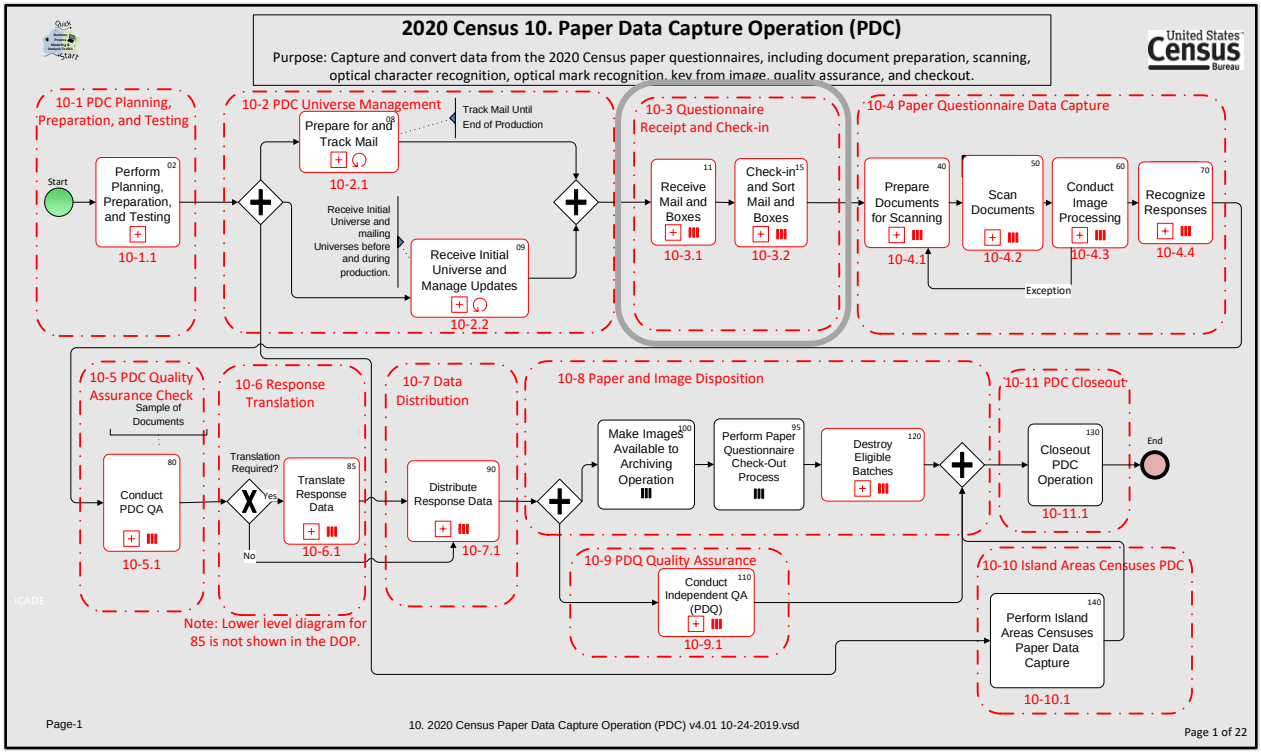

Figure 14: Questionnaire Receipt and Check-in [PDC 10-3]

The Questionnaire Receipt and Check-in activity area is subdivided into the following operational subactivities:

- Receive Mail and Boxes [PDC 10-3.1].
- Check-in and Sort Mail and Boxes [PDC 10-3.2].

Subsequent sections describe Questionnaire Receipt and Check-in operational subactivities in detail.

3.3.1 Receive Mail and Boxes [PDC 10-3.1]

A detailed view of the constituent activities that make up the "Receive Mail and Boxes" operational subactivity is given in Figure 15 below.

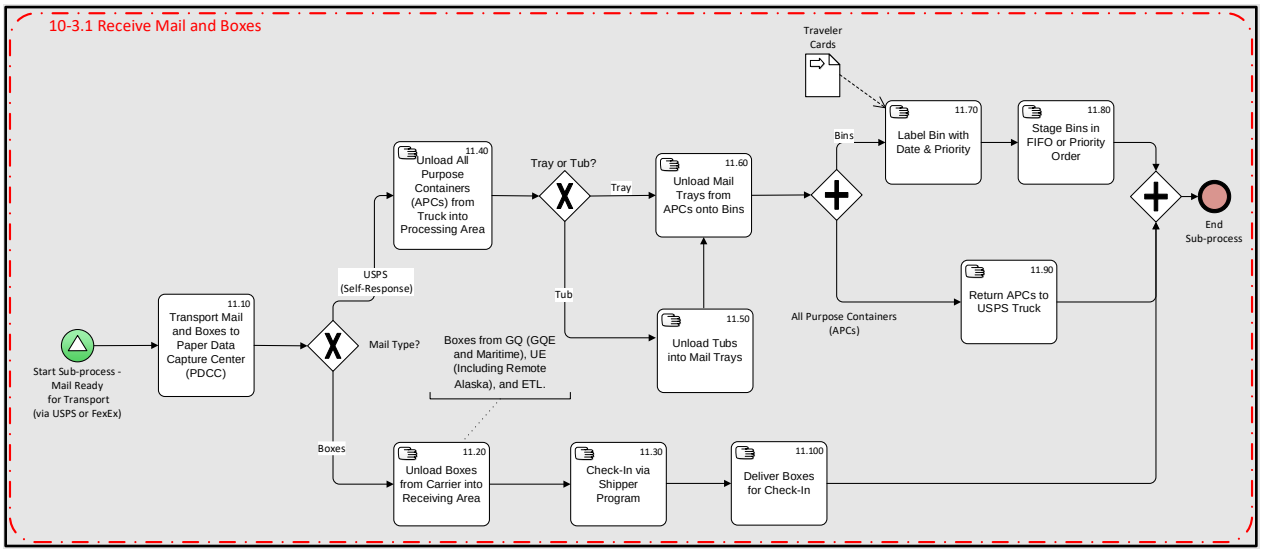

Figure 15: Receive Mail and Boxes

Receive Mail and Boxes has a path for boxes (received from the GQ, UE, and ETL operations) and another path for self-response questionnaires mailed in by respondents. General USPS mail returns are received at both PDCCs for processing. Boxed returns are delivered by commercial shipping companies to PDCC East only.

When boxed returns arrive at PDCC East, they are unloaded from the USPS delivery vehicle and carried into the receiving area. After initial verification of receipt, the boxes are delivered for Check-in.

For self-response questionnaires mailed in by respondents, the USPS uses tractor-trailers to transport and deliver forms coming from their local distribution centers to the PDC facilities. (In some cases, smaller trucks deliver mail from secondary sites.) Each tractor-trailer holds 36 USPS All Purpose Containers (APCs). Containers are unloaded from the truck into the processing area. Mail trays and mail tubs are offloaded from APCs, sorted, and then placed on bins that are labeled with a date. APCs are then returned to USPS.

Bins are staged in First-In-First-Out (FIFO) order or as required by business rules. The USPS destroys UAA questionnaires to prevent UAAs from being delivered to the PDC facilities, however, some UAAs (e.g., from experimental mail panels that did not request USPS UAA

destruction) and misdirected mailpieces are received. PDCC staff check and organize the incoming mail before automated sorting occurs. Misdirected mail is identified and set aside for return to the USPS, and unneeded UAAs are destroyed.

3.3.2 Check-in and Sort Mail and Boxes [PDC 10-3.2]

A detailed view of the constituent activities that make up the "Check-in and Sort Mail and Boxes" operational subactivity is given in Figure 16 below.

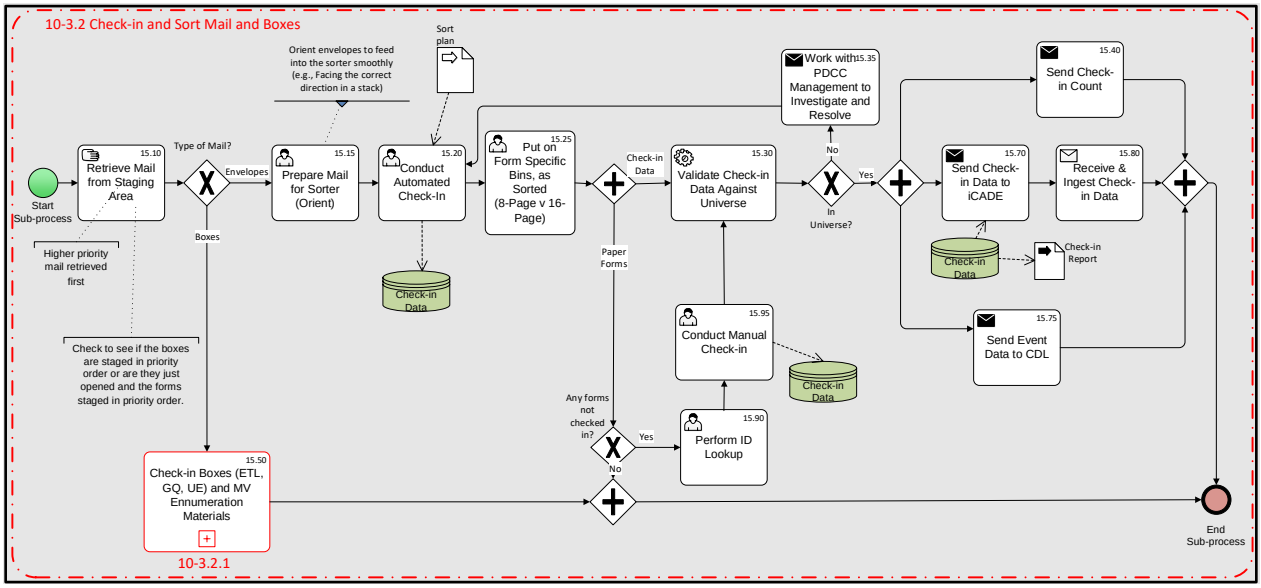

Figure 16: Check-in and Sort Mail and Boxes

The Check-in and Sort Mail and Boxes activity area consists of a single operational subactivity:

- Check-in and Sort Mail and Boxes [PDC 10-3.2].
    - Check-in Boxes (ETL, GQ, UE) and MV Enumeration Materials [PDC 10-3.2.1].

Once the USPS mail returns are received, they are placed in mail trays then sent to the sorter. There the packages are checked in using the Census ID, slit open in preparation for extraction, presorted by form type for processing, and placed back in a mail tray where they are staged for a first in-first out (FIFO) flow to envelope removal.

All receipts are processed on a FIFO basis. Forms are sorted by form type during sorter processing. PDCC West receives both 8-page English (D-Q1) and 16-page English/Spanish (D-Q1(ES)) forms, whereas PDCC East only receives the 8-page English form. The particular processing steps are:

- 2020 Census questionnaires are checked in using ATAC and receipts are sorted by a sorter.
- The sorter separates receipts by form type according to a preestablished sort plan programmed into the sorter.
- Sorters are programmed to separate good receipts from the rejects by the ability to read the barcode visible on the outside of each response envelope.
- Any forms not checked in (rejected forms) are manually checked in.
- Data are sent to downstream systems.
- Forms are sent to document preparation.

3.3.2.1 Check-in Boxes (ETL, GQ, UE) and MV Enumeration Materials [PDC 10-3.2.1]

A detailed view of the constituent activities that make up the "Check-in Boxes (ETL, GQ, UE) and MV Enumeration Materials" operational subactivity is given in Figure 17.

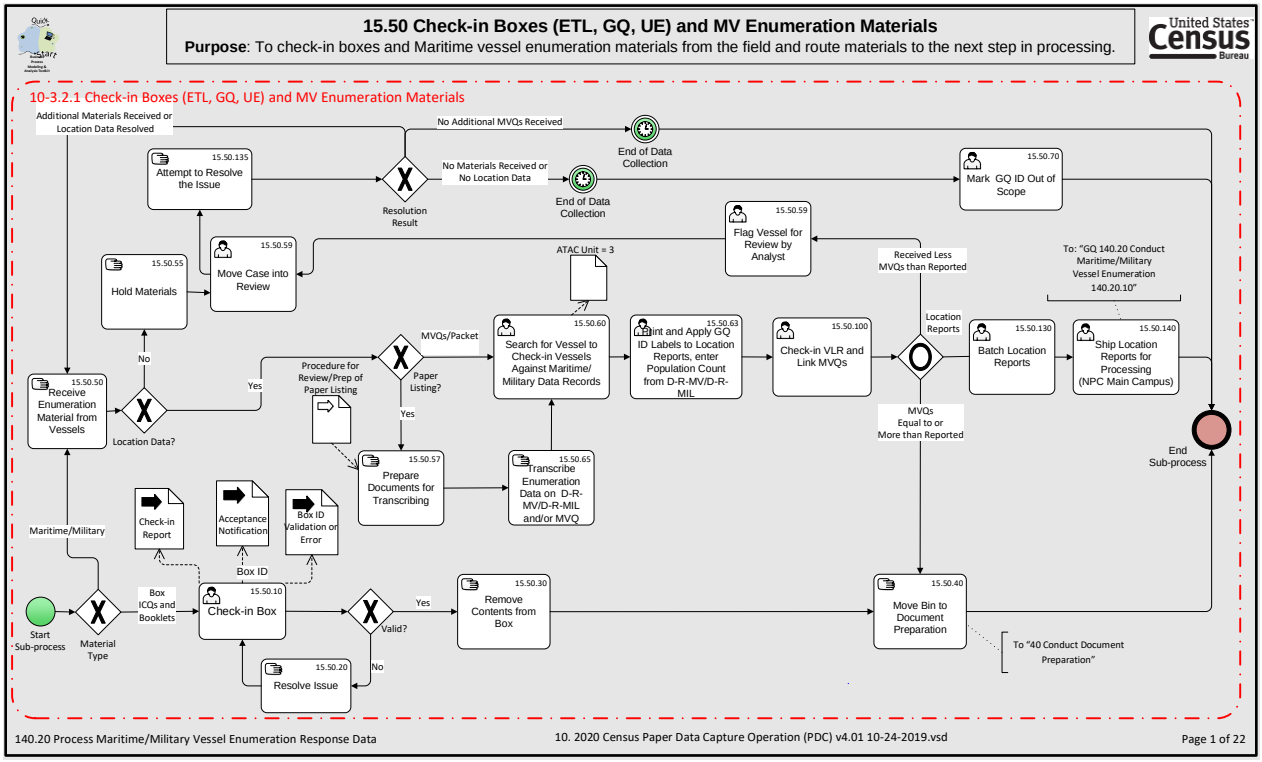

Figure 17: Check-in Boxes (ETL, GQ, UE) and MV Enumeration Materials

The subactivity Check-in Boxes (ETL, GQ, UE) and MV Enumeration Materials details the flow of boxes received at NPC. Most boxes of questionnaires take the path shown on the lower branch in the diagram for boxes of ICQs and booklets (i.e., not maritime/military vessel materials). For this path, boxes are checked in at the box level. The contents within the box are removed and

prepared for scanning. The MVQ box check-in path is separate, as shown on the upper branch of Figure 17.

Box check-in starts mid-January 2020. Box check-in occurs at PDCC East and NPC Main Campus. Boxed questionnaires are delivered to PDCC East, and boxed nonquestionnaire materials go to NPC Main Campus. If boxed and/or USPS-received questionnaires are received at NPC Main Campus, they are to be securely delivered to PDCC East. Nonquestionnaire materials delivered to PDCC East are securely delivered to NPC Main Campus.

3.4 Paper Questionnaire Data Capture [PDC 10-4]

Figure 18 below shows the BPM for the Paper Questionnaire Data Capture [PDC 10-4] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

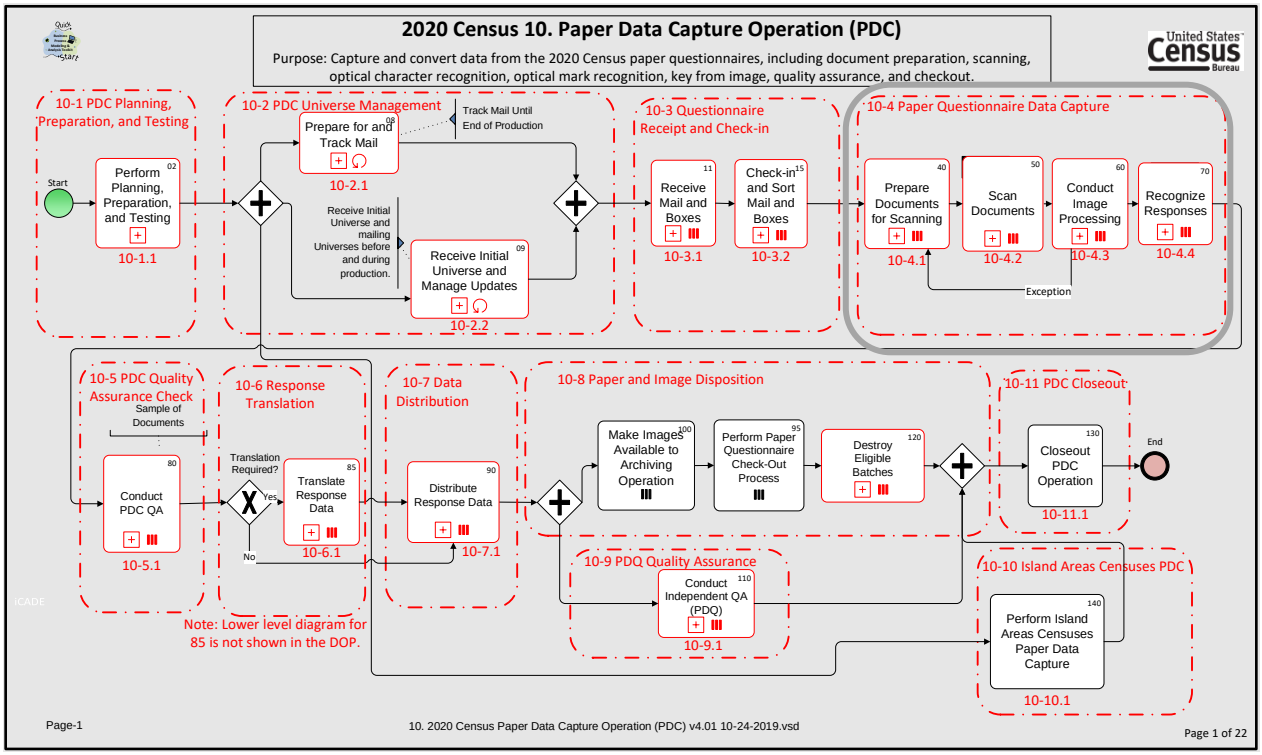

Figure 18: Paper Questionnaire Data Capture [PDC 10-4]

The Paper Questionnaire Data Capture activity area is subdivided into the following operational subactivities:

- Prepare Documents for Scanning [PDC 10-4.1].
- Scan Documents [PDC 10-4.2].

- Conduct Image Processing [PDC 10-4.3].
- Recognize Responses [PDC 10-4.4].

Subsequent sections describe the Paper Questionnaire Data Capture operational subactivities in detail.

3.4.1 Prepare Documents for Scanning [PDC 10-4.1]

A detailed view of the constituent activities that make up the "Prepare Documents for Scanning" operational subactivity is given in Figure 19 below.

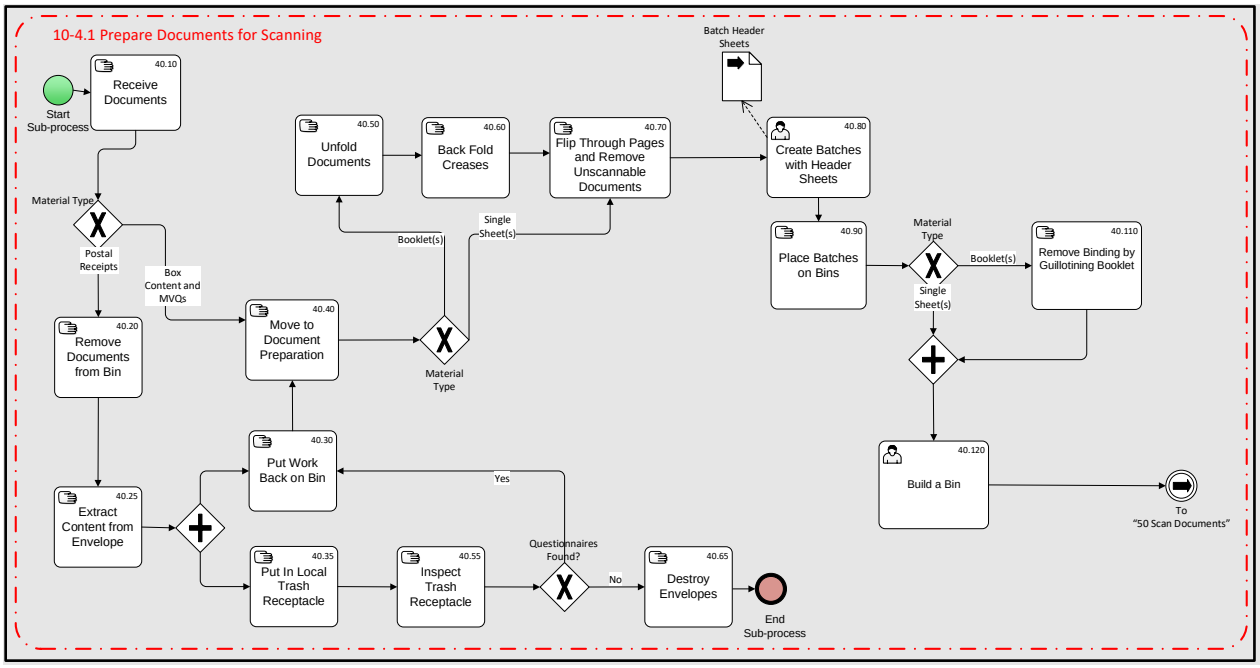

Figure 19: Prepare Documents for Scanning

Prepare Receipts - Document Preparation:

- Booklets returned in an envelope are removed from the envelope and unfolded. Boxed booklets and single sheet GQ forms should already be unfolded.
- Forms are assembled into batches and prepared for scanning.

Once forms are at envelope removal for processing, a clerk removes the forms from the envelope and discards the envelopes in cardboard bins under their table. The cardboard bins are numbered to correspond with workstations for accountability during later inspection.

Once removed from their envelopes, the forms are placed inside a different mail tray. Once a mail tray is full, conveyor belts/rollers move the full mail trays of forms to the supervisor/lead

to be placed back onto bins. The full bin of mail trays is then moved to the staging area for Doc Prep/Batching. Cardboard bins full of empty envelopes are also sent down the conveyor/roller for laborer staff to retrieve. Laborer staff inspect the boxes to help ensure no form has been inadvertently discarded. The empty envelopes are put into a gaylord (large cardboard box) and sent to NPC Main Campus for shredding.

The mail trays of forms move to batching, where batches are created and associated with a batch header sheet. A batch is a group of forms containing 500 sheets or less. Forms are not split across batches (all pages of a particular form are kept together in a batch). The barcode of each questionnaire is manually scanned, and the form is associated with a batch. Batches of single-sheet forms are sent to scanning. Batches of booklet forms are sent to have their binding removed at the guillotine to prepare them for scanning.

Build-A-Bin is conducted using a separate computer application designed to determine the flow of batched forms through iCADE. The Build-A-Bin process allows intermittent tracking of paper forms after guillotining and then again at the exception review phase, as needed.

3.4.2 Scan Documents [PDC 10-4.2]

A detailed view of the constituent activities that make up the “Scan Documents” operational subactivity is given in Figure 20 below.

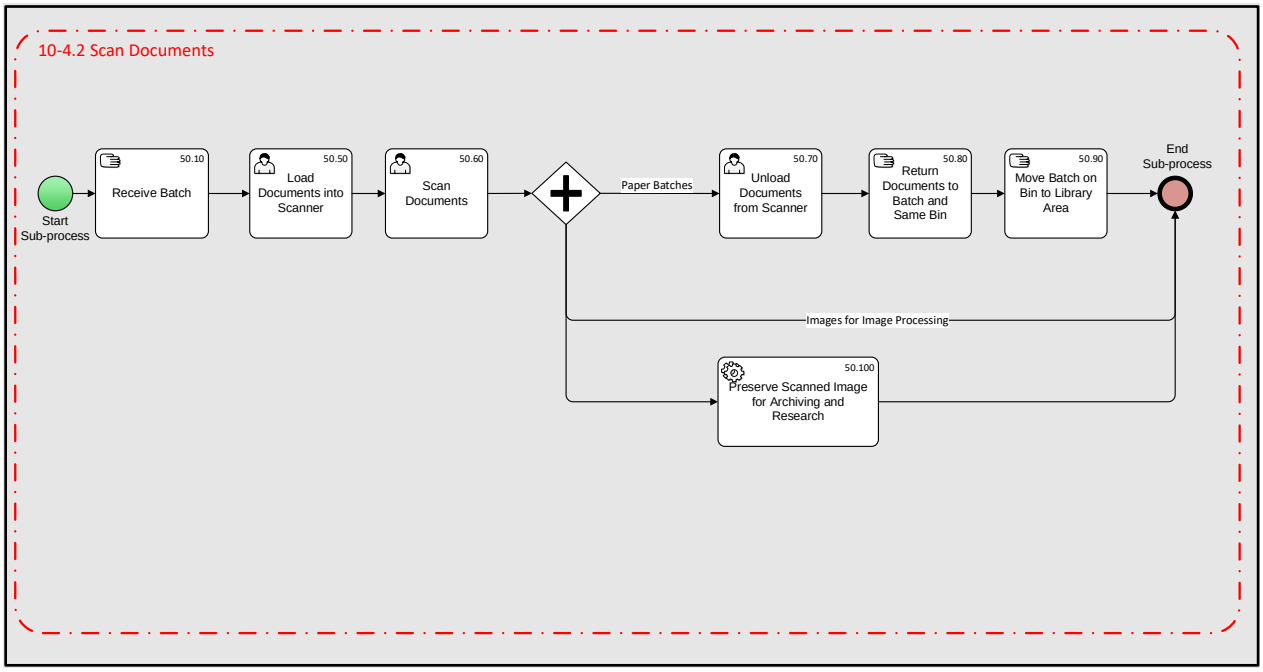

Figure 20: Scan Documents

There are a number of steps in the iCADE process of ensuring the data on the form are captured accurately on the image. Scanning personnel are responsible for scanning every page of every form within the batch. Batches of forms are scanned, utilizing an Imaging Business Machines, LLC (ibml) ImageTrac scanner, and an electronic form image is created for each form. Double feeds, scanner jams, or problem forms are rescanned as needed. Routine scanner cleaning and calibration are performed regularly by full-time Decennial maintenance support staff. After all forms are scanned, they are returned to the same bin and moved to the library area. They remain there in the event they are needed for exception review.

Scanned form images that pass certain quality criteria are sent on to the data capture process.

3.4.3 Conduct Image Processing [PDC 10-4.3]

A detailed view of the constituent activities that make up the "Conduct Image Processing" operational subactivity is given in Figure 21 below.

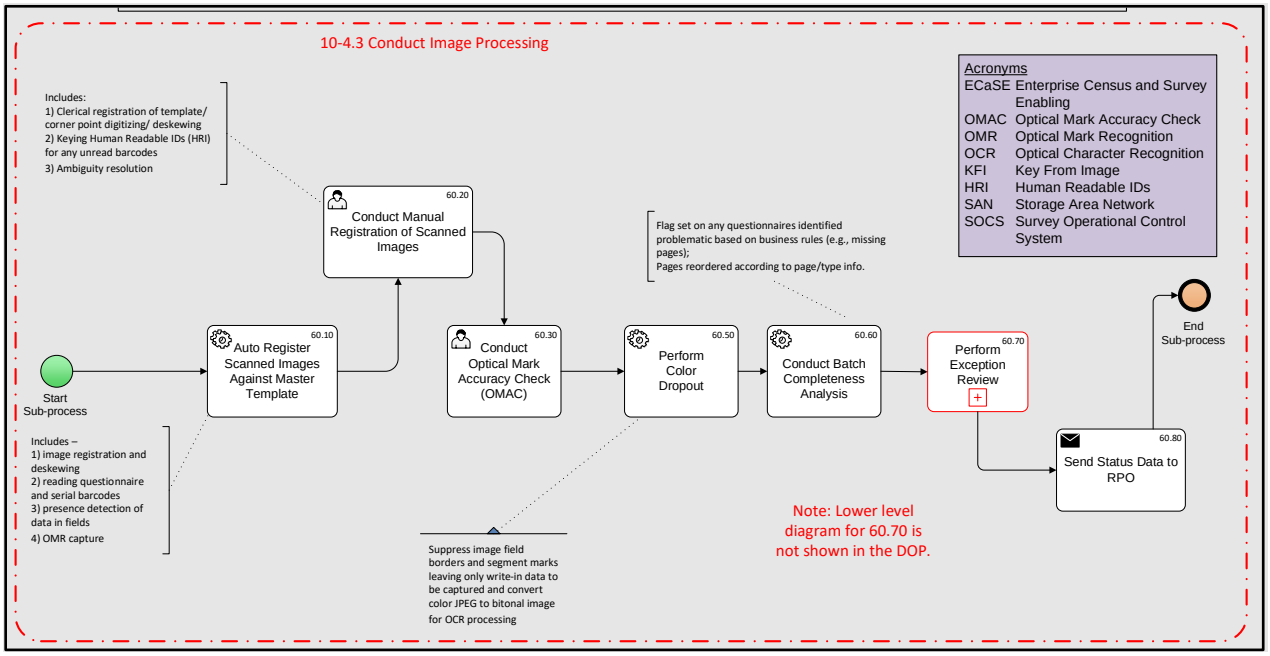

Figure 21: Conduct Image Processing

Auto Registration – After scanning, images are created and matched against the master template for the form type and page identification. This allows the system to find duplicate pages, missing pages, and unexpected forms. Next, the OMR checkboxes are read, evaluated, and captured. Problems are sent to manual registration for resolution.

Manual Registration (MR) – The iCADE Manual Registration software consists of multiple processes:

- Manual Registration of Form Page: Allows clerical repair of form images that the system did not recognize automatically. A user may not see this step if all pages registered automatically. The Manual Registration operation allows the user to repair three types of page registration failures:
  - Unread page barcodes.
  - Unread label barcodes.
  - Unrecognized corner points, which lead to a failure to link a questionnaire page with the proper page template.
- Checkbox Accuracy Review (also referred to as "Noise" or Dirty Checkbox Review): Enables a quality check of all checkbox fields read automatically by the OMR software.
- Checkbox Ambiguity Repair: Enables clerical resolution of checkbox marks identified by the Optical Mark Recognition (OMR) software as either ambiguous or not in accordance with business rules.
- OMAC (Optical Mark Accuracy Check): Quality Assurance check of the work performed by the initial manual registration clerk. This QA process has two steps:
  - OMAC Verification.
  - OMAC Adjudication.

Batch Completeness Analysis – This step ensures that all form IDs that were expected are present and accounted for, that all pages within a form are present and in sequence, and that there are no issues such as missing pages, extra pages, duplicate pages, missing form IDs, extra form IDs, page count issues, or duplicate form IDs. Any issues identified during this processing step are flagged for manual examination at Exception Review (ER).

Exception Review – This process provides an opportunity to repair problems identified after the scanning operation and before the images proceed to keying. The MR process matches the questionnaires scanned with the list of questionnaires recorded at batching and identifies those questionnaires batched but not seen at scanning (or scanned but not recorded at batching). The ER process then examines the page ID barcode for each image coming from the scanner to identify missing pages or duplicate pages. The software allows clerks to pull questionnaires and send to repair or mark as "keep" depending on operation-developed criteria. If no problems exist within a batch, the system bypasses the ER process altogether and the batch is sent automatically to KFI.

3.4.4 Recognize Responses [PDC 10-4.4]

A detailed view of the constituent activities that make up the "Recognize Responses" operational subactivity is given in Figure 22 below.

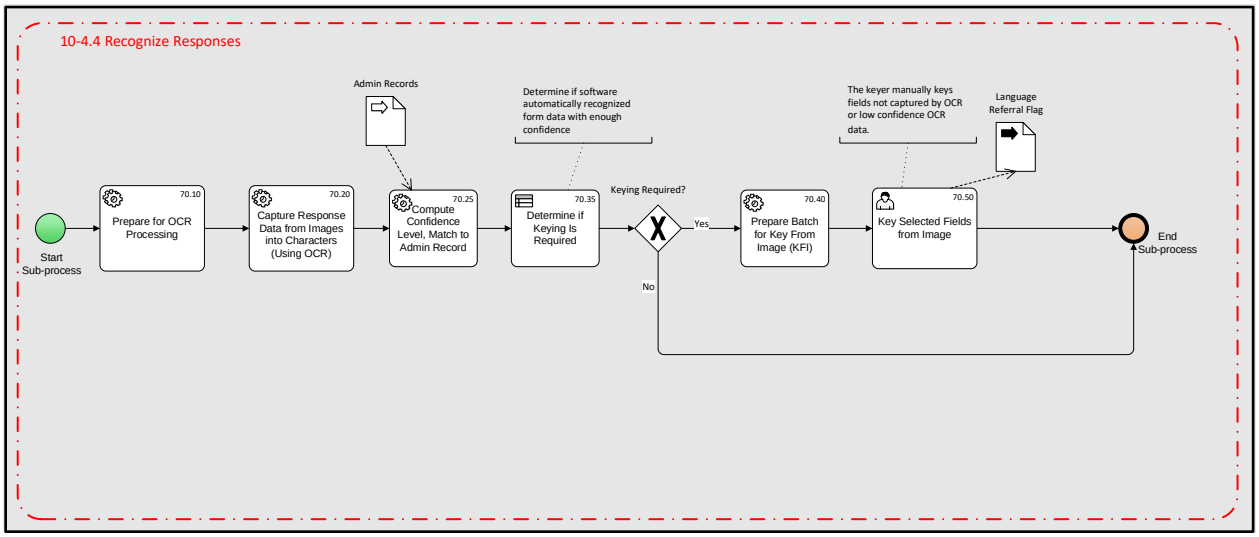

Figure 22: Recognize Responses

A template for each form type has been incorporated into the system. The software allows the definition of all the OCR (and OMR) zone areas that are used to recognize data contained on the form. An input image is registered, and the OCR (and OMR) zones are adjusted to maximize the optical recognition processing accuracy. OMR is used to read responses that are in check boxes; OCR is used to read alphabetic and numeric characters. Data are captured in English and Spanish only.

The system uses field-level confidence thresholds to accept or reject characters written in the field. In cases where the recognition software cannot automatically recognize data from the form image, the image is sent to manual capture where a keyer performs Key From Image (KFI).

The purposes of KFI are to capture field-level data manually that could not be captured automatically at the required quality level and to process QA samples. The KFI step happens after scanning and character recognition have been performed.

The keyer uses a computer workstation to view the field of a questionnaire that has been presented for review and correctional keying. The information keyed into the system should be identical to what is on the form or as per capture and business rules. For example, the KFI specification contains keying rules that instruct keyers to ignore some characters such as special characters and alphabetic characters found in the Date of Birth write-in field.

3.5 PDC Quality Assurance Check [PDC 10-5]

Figure 23 below shows the BPM for the PDC Quality Assurance Check [PDC 10-5] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

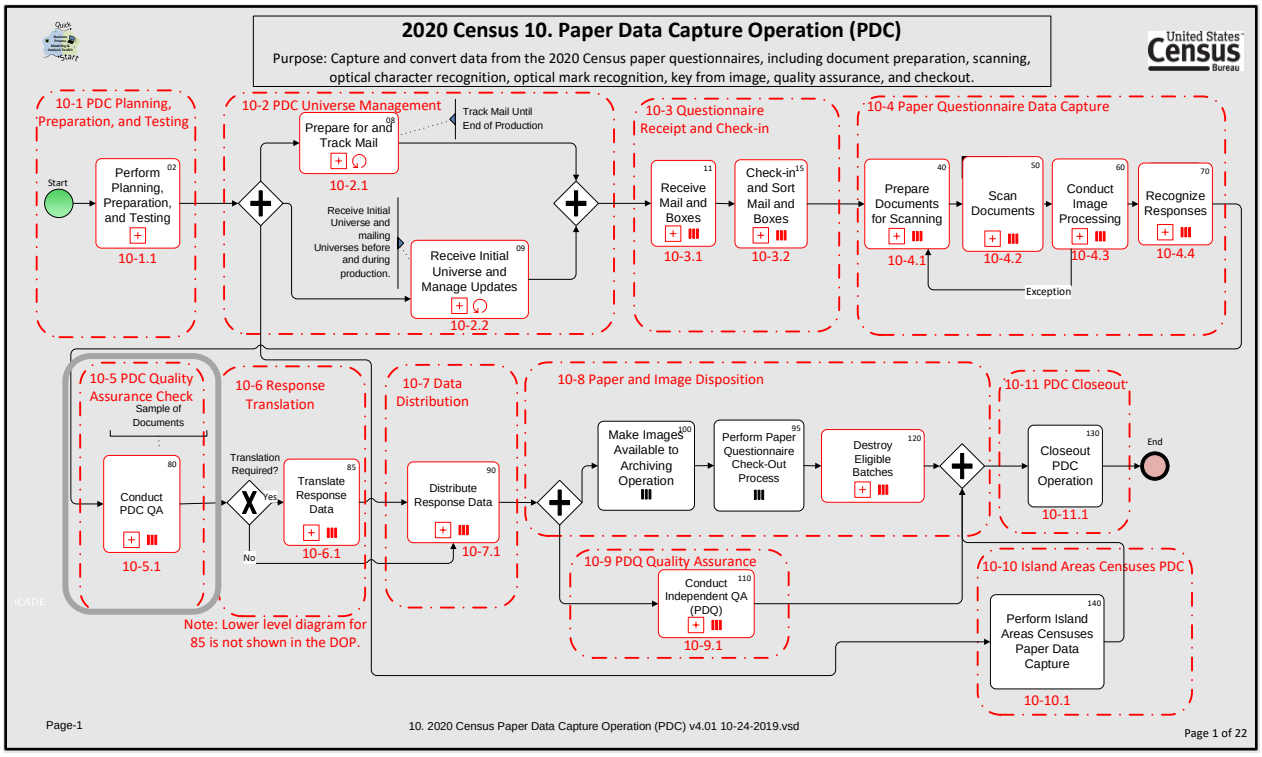

Figure 23: PDC Quality Assurance Check [PDC 10-5]

The PDC Quality Assurance (QA) Check activity area consists of a single operational subactivity.

- Conduct PDC QA [PDC 10-5.1].

The next section describes the PDC Quality Assurance Check operational subactivity in detail.

3.5.1 Conduct PDC QA [PDC 10-5.1]

A detailed view of the constituent activities that make up the "Conduct PDC QA" operational subactivity is given in Figure 24 below.

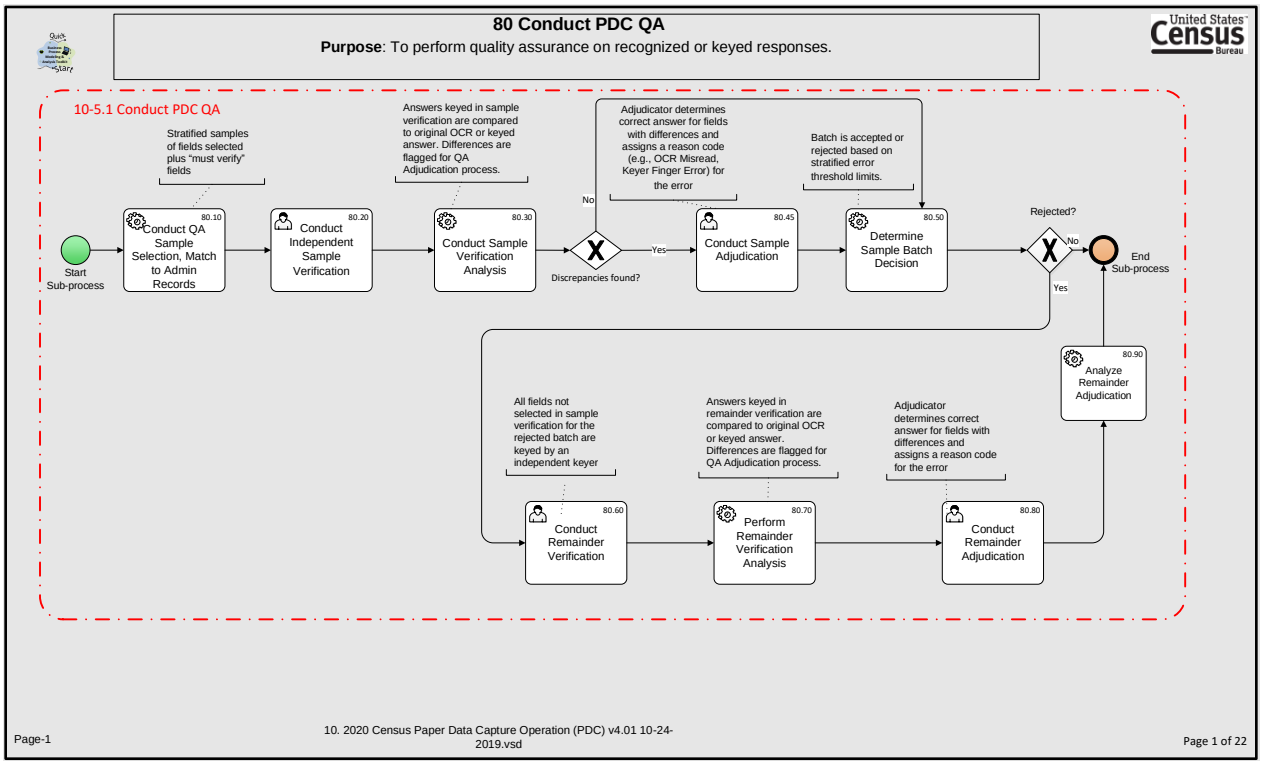

Figure 24: Conduct PDC QA

A random sample of OCR Nonblank and KFI Nonblank fields is selected for sample verification. Answers keyed in the sample verification are compared with OCR or keyed responses, and differences are flagged for the QA classification process. The classifier determines the correct answer fields with differences and assigns reason codes for the error. The batch is accepted or rejected based on error threshold limits. The target quality levels used are as follows: OMR, 99 percent; OCR, 97 percent; KFI, 99 percent. (These are the same target quality levels used during the 2010 Census.)

The universes of OCR captured fields and KFI captured fields are accepted or rejected separately. If the batch passes both OCR and KFI, it is sent to Output. Batches that are rejected only have strata fields that fail sent back for rekeying, an action that is called Remainder Processing. If it is determined that the batch requires Remainder Processing, all fields not selected in the original verification sample are presented to a keyer different from the original keyer for remainder processing. The keyer for KFI Remainder Verification is not able to see what the original keyer entered. The answers keyed for the selected fields are keyed and then compared against the original keyer's data.

All keyed batches are tracked by keyer daily through a series of iCADE reports and based on sponsor criteria. If a keyer's error rates are high, their qualification can be impacted, and the keyer may ultimately be removed from the data capture process. The system computes differences in all the fields, the types of differences, and the error rates.

3.6 Response Translation [PDC 10-6]

Figure 25 below shows the BPM for the Response Translation [PDC 10-6] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

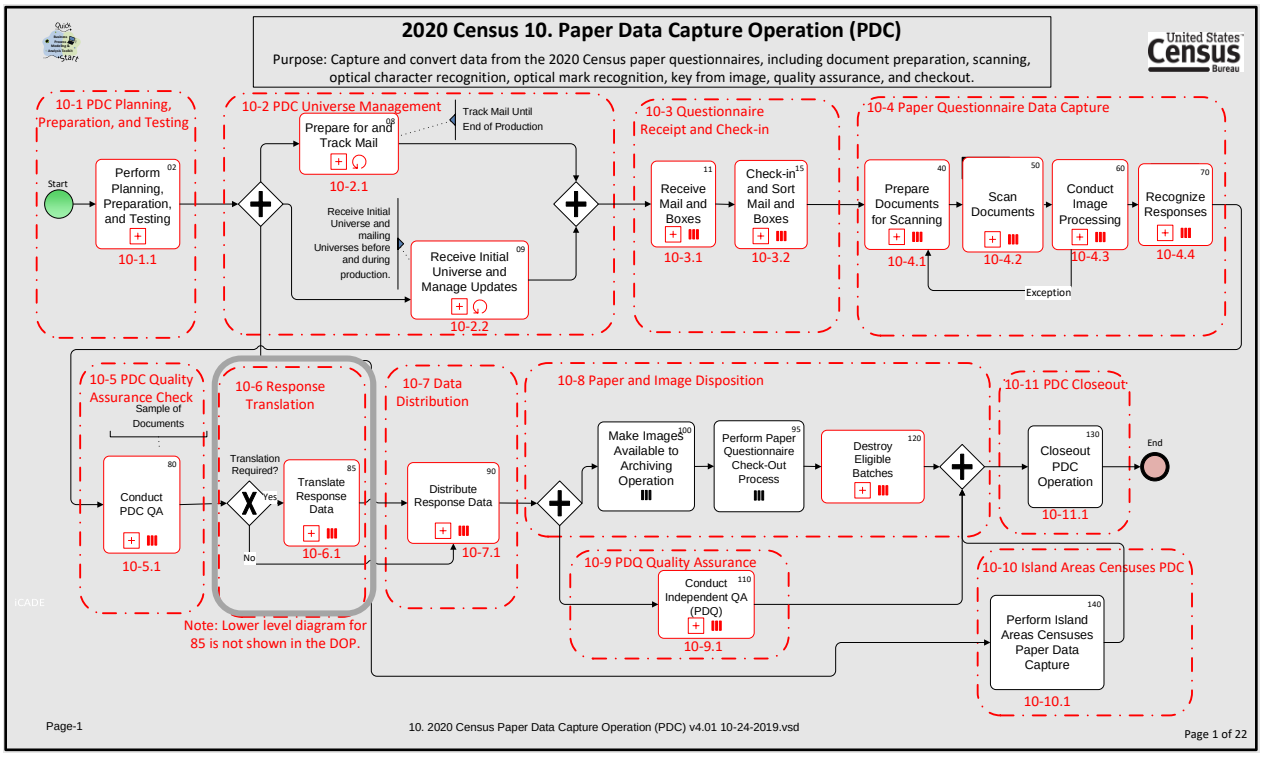

Figure 25: Response Translation [PDC 10-6]

The Response Translation activity area consists of a single operational subactivity.

- Translate Response Data [PDC 10-6.1].

The actions for Translate Response Data are:

- Low confidence responses or responses submitted in another language other than English or Spanish are flagged and sent to manual key entry.

- Manual key entry clerks determine illegible response data from non-English and non-Spanish responses.
- Clerks key response data provided in Spanish. All other non-English languages are flagged.
- Once a batch is completed and if any Language Referral flags were set, that batch now flows to the translation key entry clerks.
- Once all the flagged translation fields are processed, the batch moves to the next phase in processing.

The next section describes the Translate Response Data subactivity in detail.

3.6.1 Translate Response Data [PDC 10-6.1]

The "Translate Response Data" operational subactivity is represented by a single process step shown in the gray rounded rectangle in Figure 25 above.

For the 2020 Census, PDC is required to translate certain fields on questionnaires that are written in a language other than English or Spanish. This plan was tested in the 2018 End-to-End Census Test (CT).

As paper forms are scanned, questionnaires completed in another language fail OCR. Failures are directed to a manual KFI operation. Keyers then view an image and determine if the handwriting is in another language. The keyer flags fields that are filled out in a language other than English or Spanish. Specified translation-eligible fields are held for translation by Tucson Contact Center (TCC) staff. The TCC staff access iCADE, view the flagged fields, then translate and key responses into the iCADE system in English.

3.7 Data Distribution [PDC 10-7]

Figure 26 below shows the BPM for the Data Distribution [PDC 10-7] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

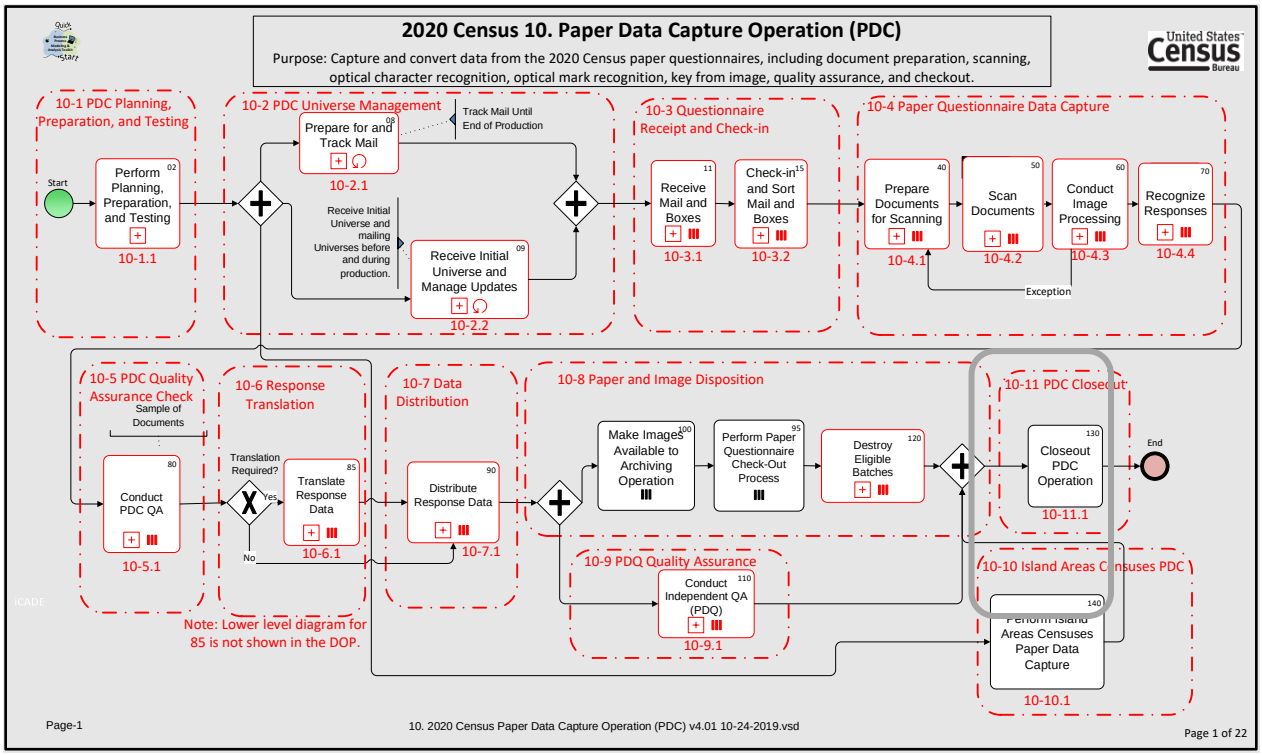

Figure 26: Data Distribution [PDC 10-7]

The Data Distribution activity area consists of a single operational subactivity.

- Distribute Response Data [PDC 10-7.1].

The next section describes the Data Distribution subactivity in detail.

3.7.1 Distribute Response Data [PDC 10-7.1]

A detailed view of the constituent activities that make up the “Distribute Response Data” operational subactivity is given in Figure 27 below.

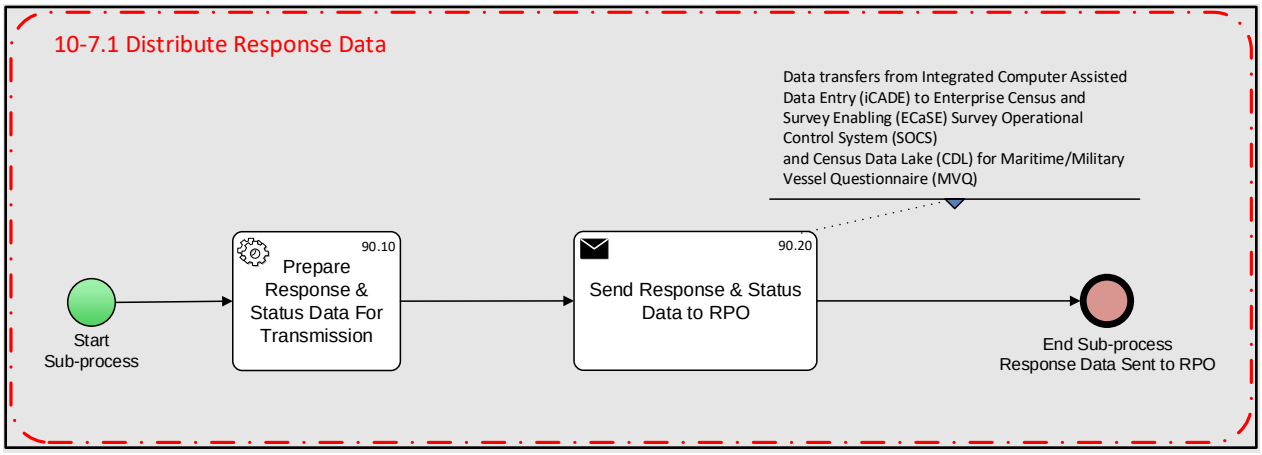

Figure 27: Distribute Response Data

Send Response Data to Response Processing Operation

Once completed, PDC delivers completed response data in multiple intervals during the process to the Decennial data delivery interface. The file delivery process provides the PDC operation confirmation that the file was received by RPO. The accepted file delivery is just one of the criteria in the approval to destroy the paper form.

Send Case Status Data to Response Processing Operation

The iCADE system sends status data to various areas several times per day. The information provided includes processing status scanning, keying, and response data that is sent to areas such as Enterprise Census and Survey Enabling (ECaSE) and CDL. The within-system iCADE reports show work moving through stages in their capture process.

3.8 Paper and Image Disposition [PDC 10-8]

Figure 28 below shows the BPM for the Paper and Image Disposition [PDC 10-8] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

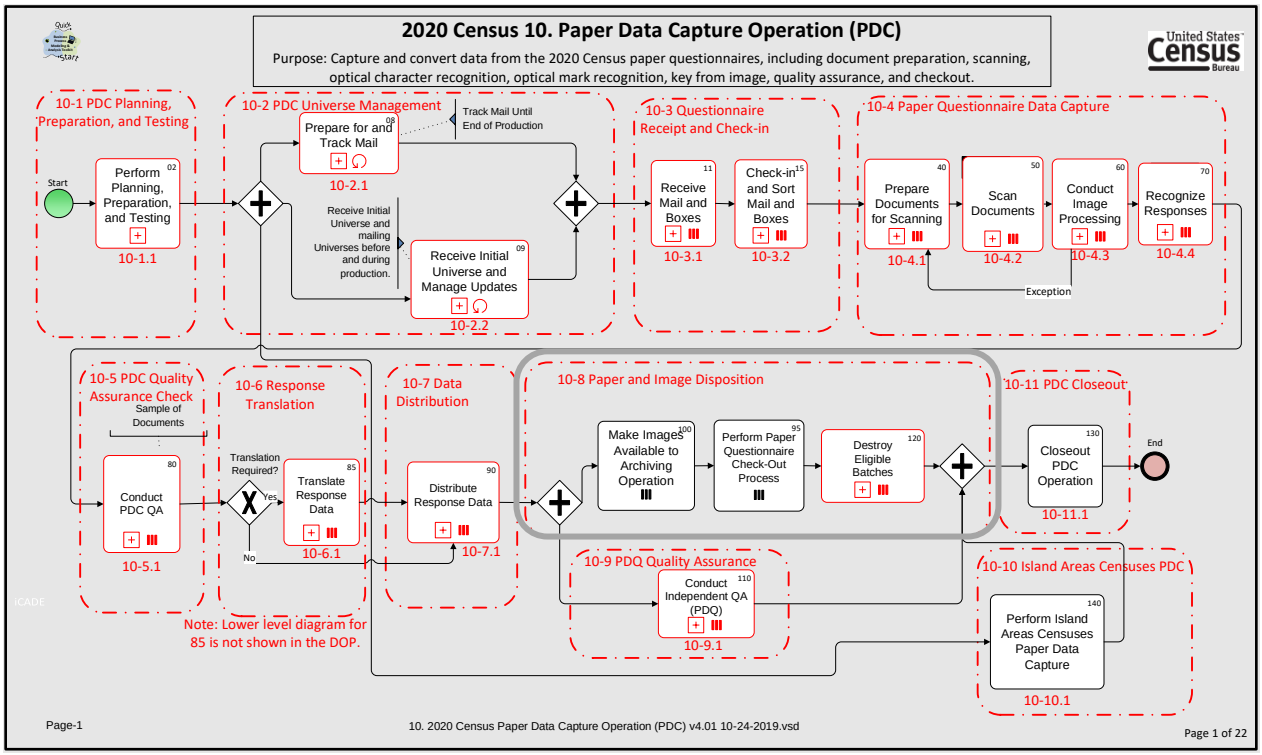

Figure 28: Paper and Image Disposition [PDC 10-8]

The "Paper and Image Disposition" operational subactivity is represented by the three process steps shown in the gray rounded rectangle in Figure 28 above.

The three steps listed are: Make Images Available to Archiving Operation, Perform Paper Questionnaire Check-out process, and Destroy Eligible Forms.

Form Checkout – The forms checkout function is the last step in the data process flow. It occurs before the forms in each batch can be sent for destruction. This function confirms that all forms were removed from the batch in Exception Review and have been reprocessed through the applicable processes, and that response data and images for the batch have been successfully transmitted and backed up.

3.9 PDQ Quality Assurance [PDC 10-9]

Figure 29 below shows the BPM for the Paper Data Quality (PDQ) Quality Assurance [PDC 10-9] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

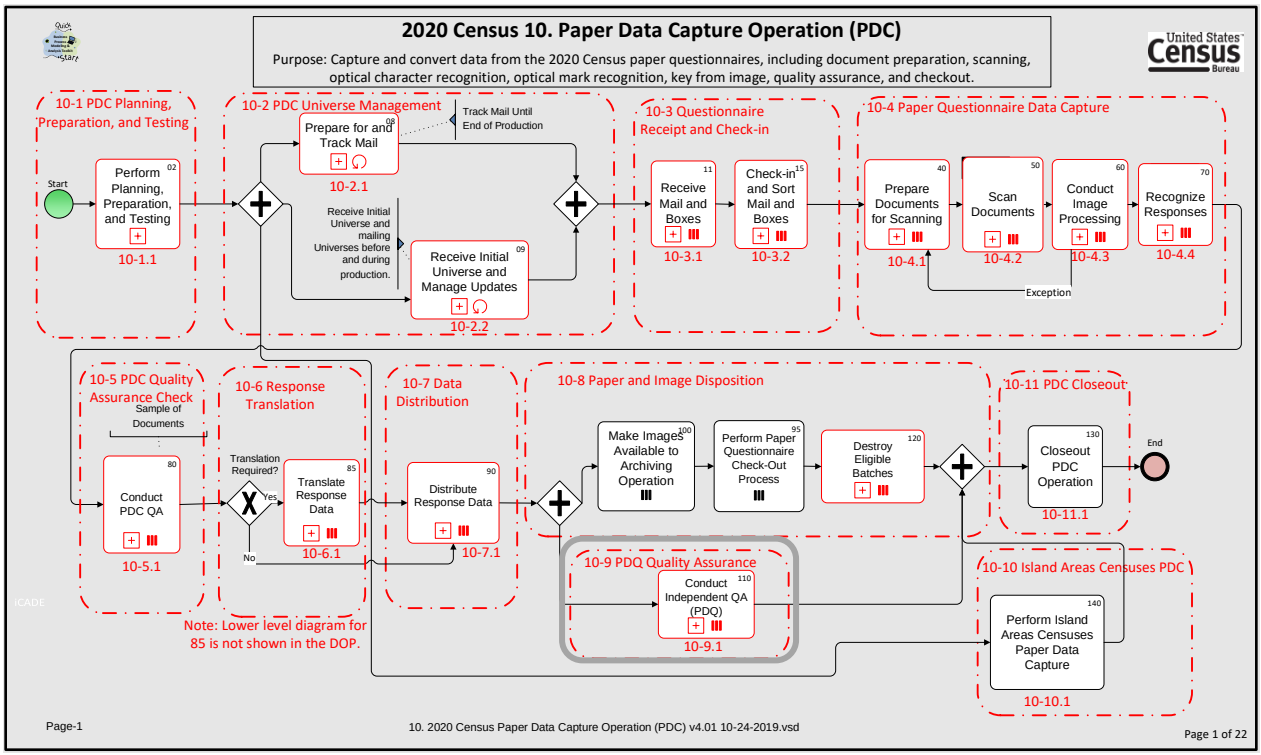

Figure 29: PDQ Quality Assurance [PDC 10-9]

The PDQ Quality Assurance activity area consists of a single operational subactivity.

- Conduct Independent QA (PDQ) [PDC 10-9.1].

The next section describes the PDQ Quality Assurance subactivity in detail.

3.9.1 Conduct Independent QA (PDQ) [PDC 10-9.1]

A detailed view of the constituent activities that make up the “Conduct Independent QA (PDQ)” operational subactivity is given in Figure 30 below.

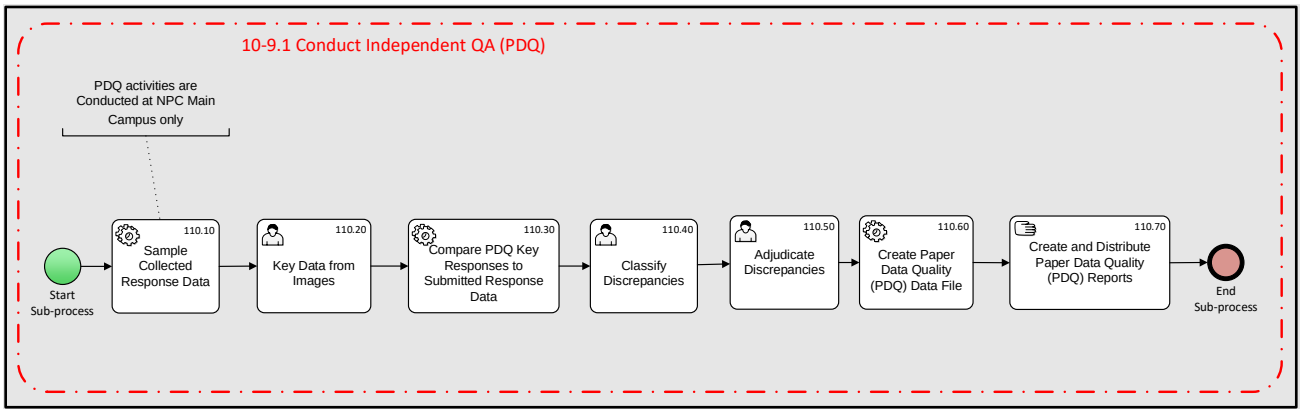

Figure 30: Conduct Independent QA (PDQ)

For the 2020 Decennial Census, a post-capture PDQ process is conducted on a sample of all form types processed through the iCADE system.

- The PDQ step is conducted on data that have already been processed and output during production.
- PDQ error coding is used as feedback to PDCC management to bring attention to any production issues.
- Due to recruiting and hiring issues, PDQ is performed after production closeout.

PDQ is conducted using Citrix for both PDCCs. The overall 2020 PDQ sampling rate is set to 1 percent of the total paper receipts.

3.10 Island Areas Censuses PDC [PDC 10-10]

Figure 31 below shows the BPM for the Island Areas Censuses PDC [PDC 10-10] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

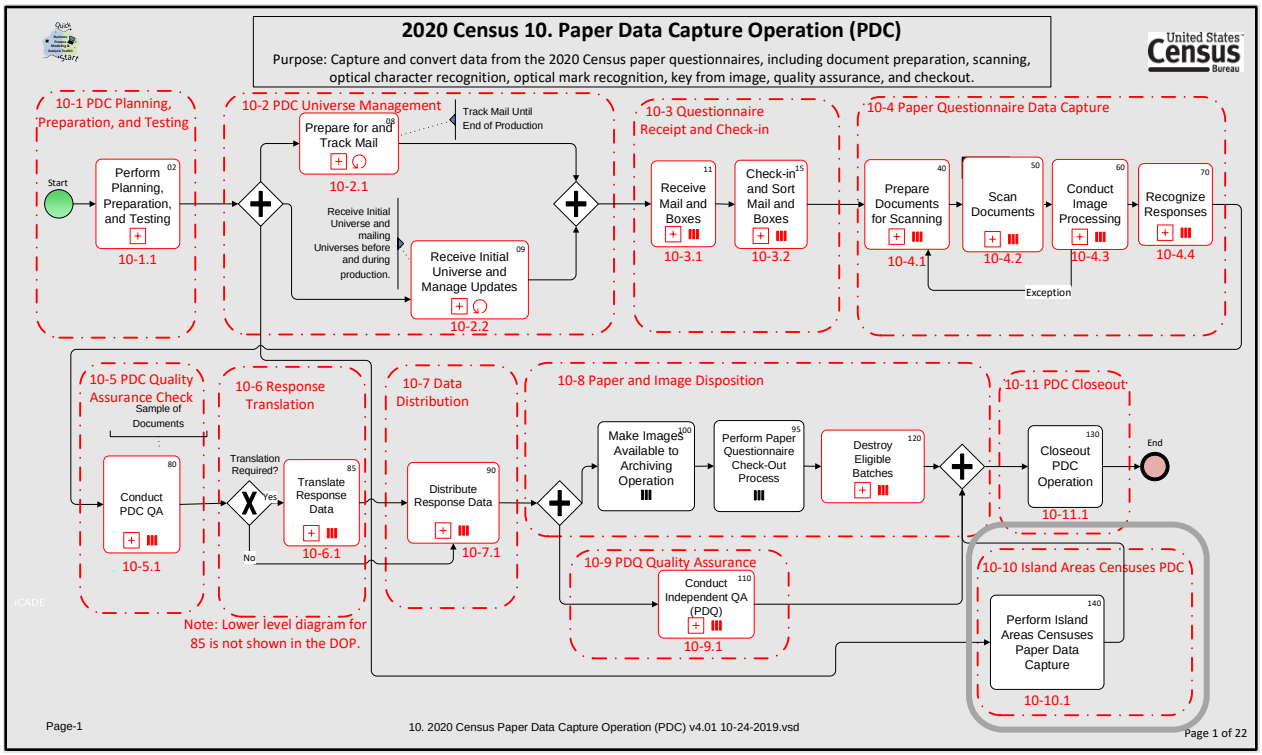

Figure 31: Island Areas Censuses PDC [PDC 10-10]

The Island Areas Censuses PDC activity area consists of a single operational subactivity.

- Perform Island Areas Censuses Paper Data Capture [PDC 10-10.1].

The next section describes the Island Areas Censuses PDC subactivity in detail.

3.10.1 Perform Island Areas Censuses Paper Data Capture [PDC 10-10.1]

The "Perform Island Areas Censuses Paper Data Capture" operational subactivity is represented by a single process step shown in the gray rounded rectangle in Figure 31 above.

The IA paper capture process follows the American Community Survey (ACS) flow for the most part. IAC PDC is conducted at NPC Main Campus. The IAC paper questionnaires are processed using the NPC current surveys platform. PDC sends the RPO response and status information related to the Island Areas Censuses (IAC) data capture work. PDC also provides image data for IAC questionnaires to the Archiving operation.

3.11 PDC Closeout [PDC 10-11]

Figure 32 below shows the BPM for the PDC Closeout [PDC 10-11] activity area (area within the gray rounded rectangle) and its constituent activities within the overall context of the PDC operation.

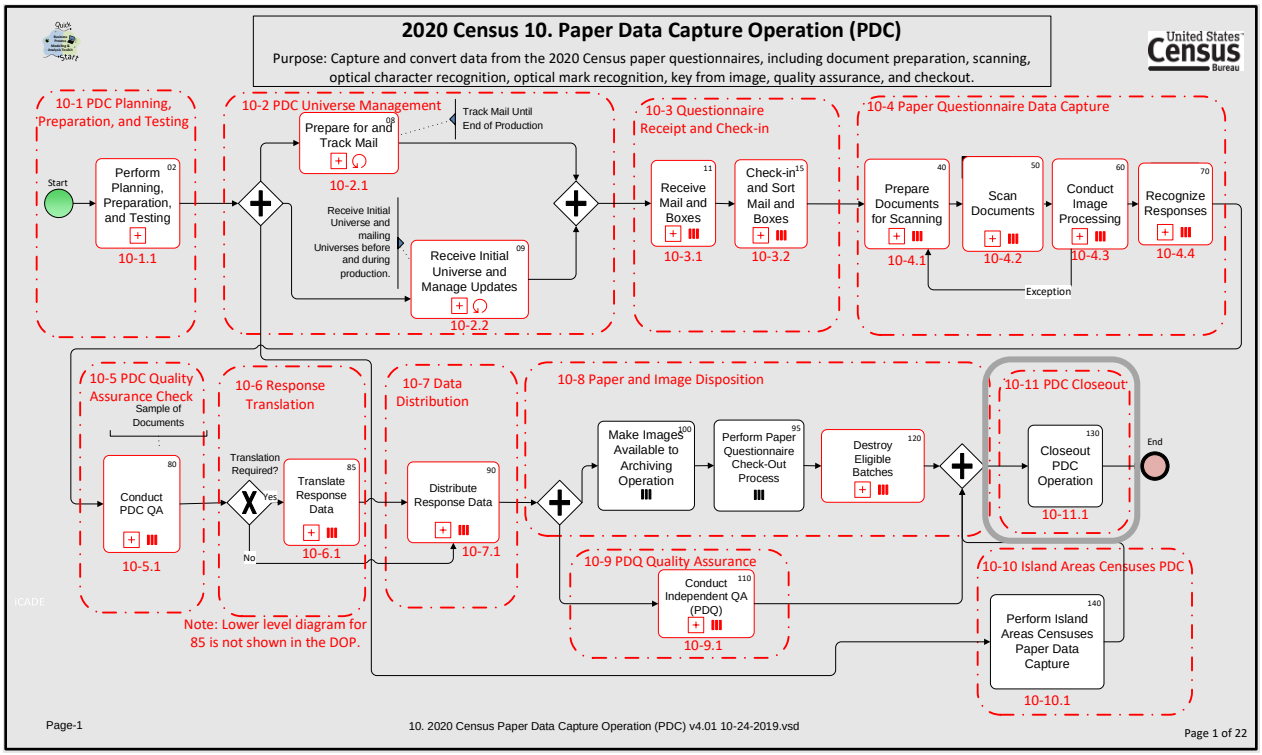

Figure 32: PDC Closeout [PDC 10-11]

The PDC Closeout activity area consists of a single operational subactivity.

- Closeout PDC Operation [PDC 10-11.1].

The next section describes the PDC Closeout subactivity in detail.

3.11.1 Closeout PDC Operation [PDC 10-11.1]

The "Closeout PDC Operation" operational subactivity is represented by a single process in the gray rounded rectangle in Figure 32 above.

PDC closeout activities include the following:

- Complete all paper processing, resolve and/or document all uncompleted cases.
- Coordinate close-out activities for leased facilities.

- Archive PDCC information/data prior to de-installation and sanitizing of systems.
- Decommission and remove all FITd equipment.
- Decommission and remove all IT, TCO, Computer Services Division equipment.
- Remove remaining Census Bureau items.
- Destroy and/or archive all 2020 Census materials.
- Develop lessons learned and other documentation for historical use.
- Perform a gradual, controlled drawdown of staff.

4. Cost Factors

4.1 Background

The investment in Paper Data Capture operation (PDC) is projected to influence the 2020 Census overall costs as described in the rest of this section.

4.2 Cost Factors

The investment in PDC is projected to influence (reduce ↓ or increase ↑) the 2020 Census overall costs in the following ways:

- ↓ Use of an enterprise solution iCADE for PDC.
- ↓ Provision of a low-cost response mode (other than the internet) to increase self-response rates.
- ↓ Use of postal tracing to reduce field operation follow-up workloads for Nonresponse Followup (NRFU).

Impacts of this operation on overall 2020 Census quality include the following:

- Plan to maintain the same quality level as the 2010 Census for OCR, OMR, and KFI.

4.3 Relevant IDEF0 Mechanisms

The following mechanisms from the IDEF0 Context Diagram represent the resources used to support this operation and will therefore impact its cost:

<u>Staff</u>

- Headquarters (HQ) staff
- Tucson translation staff
- NPC site staff

<u>Sites</u>

- Headquarters (HQ)
- National Processing Center (NPC) Main Campus
- Tucson and Jeffersonville Contact Centers
- Bowie Computer Center (BCC)
- Paper Data Capture Center (PDCC) East

- PDCC West

Systems

- Integrated Computer Assisted Data Entry (iCADE)
- IMb® Confirm Service Postal Tracking System version 2 (IPTSv2)
- Automated Tracking and Control system (ATAC)
- Production Environment for Administrative Record Staging Integration and Storage (PEARSIS)

Other

- Sorters
- Guillotines
- Conveyers
- Scanners
- Barcode readers
- HQ information technology (IT) infrastructure
- NPC Office IT infrastructure
- Census networks

5. Measures of Success

For the 2020 Census operations, the corresponding Measures of Success will be documented in the operational assessment study plans and final reports. The operational assessment study plan documents the criteria that will be used to define successful completion of the operation. The operational assessment report will provide results on whether the criteria were met.

In general, operational assessments report on planned to actual variances in budget, schedules, and production and training workloads. The corresponding Measures of Success (as documented in the operational assessment study plan) include variances that exceed established thresholds. See *Preparing for the 2020 Census Operational Assessment Study Plan* for the potential scope of assessment.

Types of success measures include:

- **Process Measures** that indicate how well the process works, typically including measures related to completion dates, rates, and productivity rates.
- **Cost Measures** that drive the cost of the operation and comparisons of actual costs to planned budgets. Costs can include workload as well as different types of resource costs.
- **Measures of the Quality** of the results of the operation, typically including things such as rework rates, error rates, and coverage rates.

See the corresponding operational assessment study plan and report for the Paper Data Capture Operation (PDC) for details on the measures of success.

Appendix A – Acronyms and Terminology

Table 7 lists the acronyms and abbreviations used within this Detailed Operational Plan document.

Additional Decennial terminology can be found on the Census Bureau Intranet for those with access.

Table 7: Acronyms and Abbreviations List

Acronym	Meaning
ACO	Area Census Office
ACS	American Community Survey
ADEP	Associate Director for Economic Programs
AdRec	Administrative Record
ARC	Archiving operation
ATAC	Automated Tracking and Control system
AZ	Arizona
BCC	Bowie Computing Center
BCR	Barcode Recognition
BPM	Business Process Model
BPMN	Business Process Model and Notation
CA	California
CDL	Census Data Lake
CEF	Census Edited File
CFD	Content and Forms Design operation

Acronym	Meaning
CI	Coverage Improvement
CIRA	Census Image and Retrieval Application
CNMI	Commonwealth of Northern Mariana Islands
COVID	Coronavirus Disease
CQA	Census Questionnaire Assistance operation
CRO	Count Review Operation
CT	Census Test
CUF	Census Unedited File
DCEO	Decennial Contracts and Execution Office
DCS2000	Data Capture System 2000
DLM	Decennial Logistics Management operation
DOP	Detailed Operational Plan
DPD	Data Products and Dissemination operation
DRF	Decennial Response File
DRIS	Decennial Response Integration System
DSC	Decennial Service Center operation
E2E	End-to-End
EAE	Evaluations and Experiments operation
ECaSE	Enterprise Census and Survey Enabling
ER	Exception Review
ETL	Enumeration at Transitory Locations operation

Acronym	Meaning
FIFO	First In First Out
FITd	Field IT Deployment
FLDI	Field Infrastructure operation
FOCS	Field Operation Control System
FPD	Forms Printing and Distribution operation
GDP	Geographic Data Processing
GEOP	Geographic Programs operation
GQ	Group Quarters; Group Quarters operation
GSA	Government Services Administration
HQ	Headquarters
IA	Island Areas
IAC	Island Areas Censuses operation
IA-CEF	Island Areas Censuses Census Edited File
IA-CUF	Island Areas Censuses Census Unedited File
IA-MDF	Island Areas Censuses Master Delivery File
ibml	Imaging Business Machines, LLC
iCADE	integrated Computer Assisted Data Entry
ICD	Interface Control Document
ICQ	Individual Census Questionnaire
ID	Identifier
IDEF0	Integrated Definition, Level 0

Acronym	Meaning
IE	Information Exchange
IMb®	Intelligent Mail barcode
IN	Indiana
IOD	Integrated Operations Diagram
IOE	Improving Operational Efficiency
IPT	Integrated Project Team
IPTSv2	IMb® Confirm Service Postal Tracking System
IOD	Integrated Operations Diagram
ISA	Interconnection Security Agreement
ISR	Internet Self-Response Operation
IT	Information Technology
ITIN	IT Infrastructure Operation
KFI	Key From Image
LUCA	Local Update of Census Addresses operation
MAF	Master Address File
MD	Maryland
MDF	Microdata Detail File
MR	Manual Registration
MV	Maritime/Military Vessel(s)
MVQ	Vessel Questionnaire (or Maritime/Military Vessel Questionnaire)
NA	Not Applicable

Acronym	Meaning
NARA	National Archives and Records Administration
NID	Non-ID Processing operation
Non-ID	Non-identifier; Non-identified
NPC	National Processing Center
NRFU	Nonresponse Followup operation
OCR	Optical Character Recognition
OMAC	Optical Mark Accuracy Check
OMR	Optical Mark Recognition
OTDR	Operational Test and Dry Run
P&S	Performance and Scalability
PDC	Paper Data Capture operation
PDCC	Paper Data Capture Center
PDQ	Paper Data Quality
PEARSIS	Production Environment for Administrative Records Staging, Integration and Storage
PM	Program Management operation
POP	Population Division
PR	Puerto Rico
PRR	Production Readiness Review
QA	Quality Assurance
QC	Quality Control
RA	Remote Alaska

Acronym	Meaning
RI	Reinterview
RPO	Response Processing Operation
SOCS	Survey Operational Control System
SPC	Security, Privacy, and Confidentiality operation
SR	Self-Response
TCC	Tucson Contact Center
TCO	Telecommunications Office
TEA	Type of Enumeration Area
TI	Technical Integrator
TLQ	Transitory Location Questionnaire
TRR	Test Readiness Review
UAA	Undeliverable As Addressed
UAT	User Acceptance Test
UE	Update Enumerate operation
UL	Update Leave operation
USPS	United States Postal Service
UTS	Unified Tracking System

Appendix B – References

Appendix B lists the documents or other resources used during the development of this Detailed Operational Plan document.

U.S. Census Bureau (2022), "2020 Census Operational Plan," Version 5.0, February 4, 2022.

U.S. Census Bureau (2018), "2020 Census Operational Plan," Version 4.0, December 31, 2018.

U.S. Census Bureau (2018), "Preparing for the 2020 Census Operational Assessment Study Plan," Draft, May 10, 2018.

Appendix C – Activity Tree for Paper Data Capture Operation (PDC)

This appendix presents the Activity Tree for the PDC operation. An Activity Tree uses an outline structure to reflect the decomposition of the major operational activities in the operation. Each activity is numbered according to its position in the outline. For example, for the current operation numbered "10," the first activity would be numbered 10-1. Subactivities under this activity would be numbered sequentially, starting again with the number one. For example, the first subactivity under the first activity would be numbered 10-1.1 the second subactivity as 10-1.2. The second activity would be numbered 10-2, and so on.

PDC Activity Tree:

- 10-1 PDC Planning, Preparation, and Testing
  - 10-1.1 Perform Planning, Preparation, and Testing
- 10-2 PDC Universe Management
  - 10-2.1 Prepare for and Track Mail
    - 10-2.1.1 Prepare for Mail Tracking
    - 10-2.1.2 Track Mail
  - 10-2.2 Receive Initial Universe and Manage Updates
    - 10-2.2.1 Prepare for Questionnaire Check-in
    - 10-2.2.2 Prepare for Boxed Materials Check-in
    - 10-2.2.3 Prepare for Vessel Questionnaire (MVQ) Check-in
- 10-3 Questionnaire Receipt and Check-in
  - 10-3.1 Receive Mail and Boxes
  - 10-3.2 Check-in and Sort Mail and Boxes
    - 10-3.2.1 Check-in Boxes (ETL, GQ, UE) and MV Enumeration Materials
- 10-4 Paper Questionnaire Data Capture
  - 10-4.1 Prepare Documents for Scanning
  - 10-4.2 Scan Documents
  - 10-4.3 Conduct Image Processing
  - 10-4.4 Recognize Responses
- 10-5 PDC Quality Assurance Check
  - 10-5.1 Conduct PDC QA
- 10-6 Response Translation
  - 10-6.1 Translate Response Data
- 10-7 Data Distribution
  - 10-7.1 Distribute Response Data

- 10-8 Paper and Image Disposition
- 10-9 PDQ Quality Assurance
    - 10-9.1 Conduct Independent QA (PDQ)
- 10-10 Island Areas Censuses PDC
    - 10-10.1 Perform Island Areas Censuses Paper Data Capture
- 10-11 PDC Closeout
    - 10-11.1 Closeout PDC Operation